

# 数据智能和产业创新

**马慧民**

经济学博士

上海大数据联盟常务副秘书长

上海市北高新股份有限公司副总经理

上海市欧美同学会常务理事

美国加州伯克利大学访问学者

《智能新零售》作者



马慧民是 科技部“物联网与智慧城市”重点专项核心专家。承担十余项大数据项目，发表十余篇大数据论文。



- 大数据行业专家
- 传播大数据思想
- 对接大数据资源

- ✓ 100+ 公开演讲
- ✓ 100+ 活动主持
- ✓ 100+ 培训授课



担任第一财经《创时代》、《632观察》等栏目特约主持人和特约嘉宾；还为《文汇报》、《东方财经》、《张江科技评论》等撰写大数据行业专栏。

电影『速度与激情』中的『天眼』



# 讨论：“互联网” + “大数据” 的悖论

▶ 更加自由

VS

更加不自由

- 1、获取信息；
- 2、发布信息；
- 3、个人隐私；

1993  
*On the Internet, nobody knows you're a dog.*



*"On the Internet, nobody knows you're a dog."*

2013  
*On the Internet, everybody knows you're a dog.*



# 案例1：家庭摄像头监听

## 女子从怀孕到哺乳被直播2年毫不知情 网友边看边截图

2018-09-28 07:53 新民晚报 评论 (0人参与)



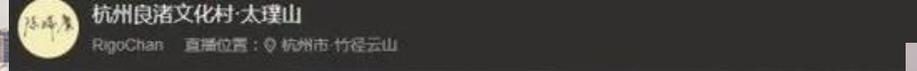
A<sup>-</sup> A<sup>+</sup>

原标题：从怀孕到哺乳，上海一女子被直播2年却毫不知情！网友边看边截图.....

如今，很多人会因为安全等原因在家里安装摄像头，可是，你知道吗？有人因为安装了摄像头，自己的一举一动全在网上直播了！

据媒体报道，上海有一位妈妈，2017年起，从其怀孕到哺乳期一直被直播。网友边观看边截图，甚至还留下极为露骨评论，这位妈妈却始终一无所知。这究竟是怎么回事呢？

据报道，近日有青果摄像头的用户反映，自己在家的生活被在网上同步直播。记者电话联系多名用户，均表示对直播不知情。记者发



案例2: 亚马逊在全球拥有一个由数千名员工组成的团队, 他们负责听取Alexa用户的语音请求, 目的是改进该软件。  
苹果Siri评估计划



数据时代, 隐私保护变得更加重要。

案例3: 2016年9月4日晚, 张强使用微信发了一句“跟我加入ISIS”。2017年6月13日, 法院判决张强犯宣扬恐怖主义、极端主义罪, 判处有期徒刑9个月, 并处罚金1000元。

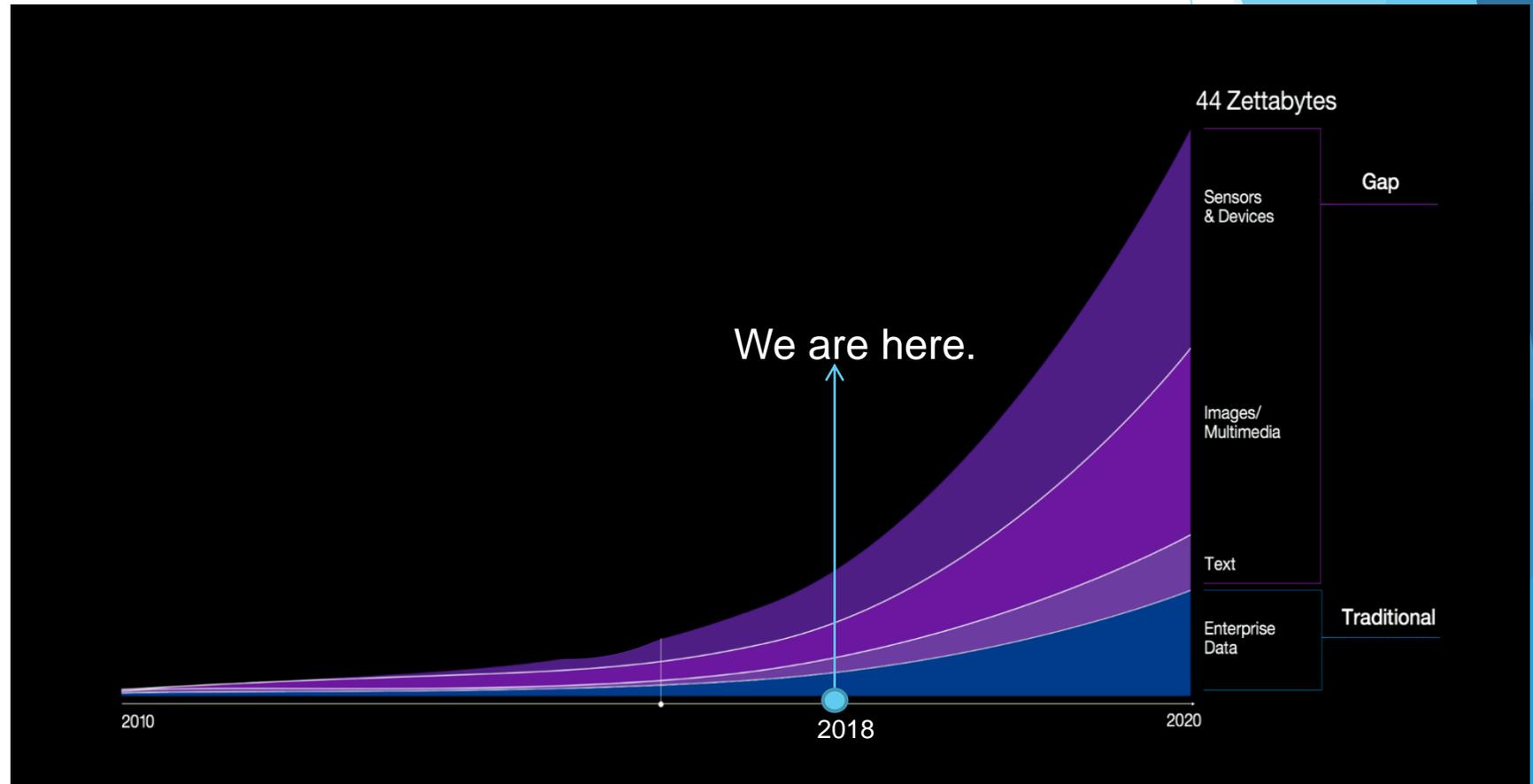


# 一、数据智能的时代趋势

# 1、什么是大数据？

**Big data** is a collection of data that can not be captured, managed and processed by conventional way within the time allowed.

4V:  
数量多(Volume);  
速度快(Velocity);  
类型多(Variety);  
价值高(Value);



$$1\text{ZB} = 10^3 \text{EB} = 10^6 \text{PB} = 10^9 \text{TB} = 10^{12} \text{GB} = 10^{15} \text{MB} = 10^{18} \text{KB} = 10^{21} \text{B}$$

大数据产业链按上中下游划分为：

资源



技术



应用

“互联网+”时代，所有企业都有潜力发展成为**数据资源公司**，处于产业链的核心地位，数据是其持续变现的资本

产业发展初期，**数据技术型公司**，包括硬件、基础软件、分析服务、信息安全等，是最直接的受益群体

数据开放的趋势下，极有可能引爆新一轮大众创业、万众创新的格局，诞生具备创新式数据思维及突出技术实力的**数据应用型公司**

广义大数据产业包含了大数据在各个领域的应用，已经超出了信息产业的范畴。

据麦肯锡预计，到2020年美国大数据应用带来的增加值将占2020年GDP的2%-4%，约为5000-8000亿美元。

对于中国市场的预测，则更加具体，根据不同领域，其市场规模也有所不同。



### 健康医疗

市场规模--预计2020年达**800亿**

年复合增长率-**20%**

主要应用--**流行病分析、预警**

### 金融领域

市场规模--预计2020年达**1300亿**

主要应用--风险管理、信誉评估、证券分析

### 交通领域

市场规模--预计2020年达**190.74亿**

年复合增长率-**100%**

主要应用--交通流量分析、预测

### 电商领域

市场规模--预计2020年达**402.57亿**；

主要应用--物流、精准营销、**CRM管理**

### 政府应用

市场规模--预计2020年达**574.47亿**；

主要应用--政府服务、数据共享、开放

# 上海大数据联盟产业云图

应用

<h3>金融</h3>	<h3>工业</h3>	<h3>能源</h3>
<h3>医疗</h3>	<h3>交通</h3>	<h3>电子商务</h3>
<h3>精准营销</h3>	<h3>人工智能</h3>	<h3>政务</h3>
<h3>其他</h3>		

技术

<h3>大数据存储与管理</h3>	<h3>大数据安全</h3>	<h3>数据分析与挖掘</h3>	<h3>大数据可视化</h3>
-------------------	----------------	------------------	-----------------

资源

<h3>基础设施</h3>	<h3>数据资源</h3>
---------------	---------------

衍生服务

<h3>咨询机构</h3>
<h3>人力资源</h3>
<h3>培训</h3>
<h3>投资</h3>

产业支撑

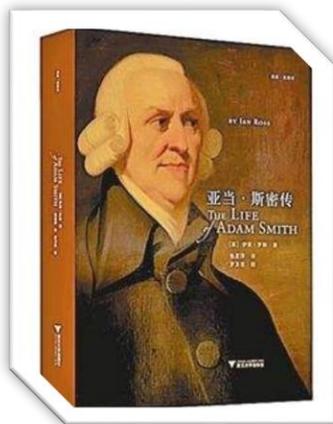
<h3>产业基地</h3>	<h3>产业联盟</h3>
<h3>交易中心</h3>	
<h3>研究机构</h3>	



# 数据成为继劳动、土地、资本之后的第四种生产要素



“劳动是财富之父，土地是财富之母”



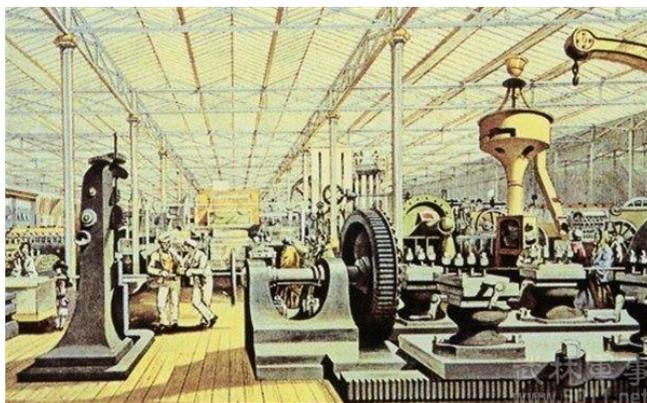
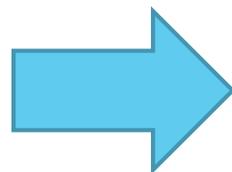
生产要素三言论：劳动、土地、资本



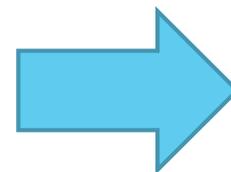
数据：虚拟空间



农业经济时代



工业经济时代



数据驱动的  
服务经济时代

## 2、大数据产生的时代背景

### 三次工业革命与六次技术革命浪潮

技术革命开始年份	该时期的流行名称	核心国家	诱发技术革命的重大技术突破	工业革命及其区间
第一次技术革命（1771）	产业革命	英国	阿克莱特在英国克隆福德设厂	第一次工业革命（1771-1875）
第二次技术革命（1829）	蒸汽和铁路时代	英国（扩散到欧洲大陆和美国）	蒸汽动力机车“火箭号”在英国利物浦到曼彻斯特的铁路上试验成功	
第三次技术革命（1875）	钢铁、电力、重工业时代	美国和德国追赶超越英国	卡内基酸性转炉钢厂在美国宾夕法尼亚州的匹兹堡开工	第二次工业革命（1875-1971）
第四次技术革命（1908）	石油、汽车和大规模生产的时代	美国（起初与德国竞争世界领导地位），后扩散到欧洲	第一辆T型车从美国密歇根州底特律的福特工厂出产	
第五次技术革命（1971）	信息和远程通讯时代（包括机器人）	美国（扩散到欧洲和亚洲）	在美国加利福尼亚州的圣克拉拉，英特尔的微处理器宣告问世	第三次工业革命（1971-2070年代？）
第六次技术革命（2020—30年代？）	新能源、3D打印、纳米、新材料、生物技术和生物电子	美国、日本和欧盟？		

# 信息时代的三个阶段：

## 计算机时代

- 从上世纪70年代末期开始，已经实现了工业化的发达国家相继开始向信息化社会转型。

## 互联网时代

- 到上世纪90年代中期，美国已经基本度过了第一阶段，计算机高度普及，解决了信息的机器可读化和数据的可计算化问题；全球主要发达国家进入信息互联时代。

## 大数据时代

- 目前，美国也基本走完了互联网时代的路程，互联网高度普及，解决了信息传递和信息服务问题。基于此，美国正在步入一个全新的历史阶段——大数据时代。

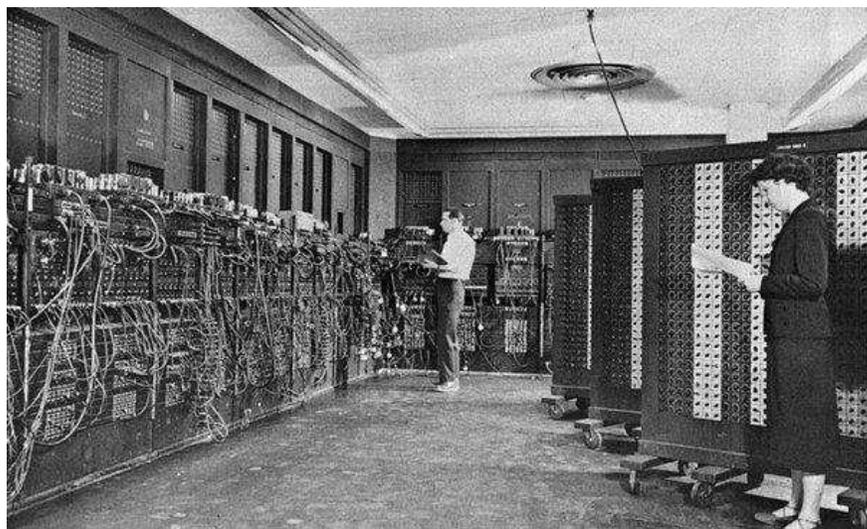
2011年5月，麦肯锡发布《大数据：创新、竞争和生产力的下一个前沿领域》。

2012年3月，奥巴马政府启动“大数据研究和发展”计划。

2012年4月，首家大数据公司Splunk上市：8-10美元——> 35.48美元，30亿美元。

2012年被称为“大数据元年”。

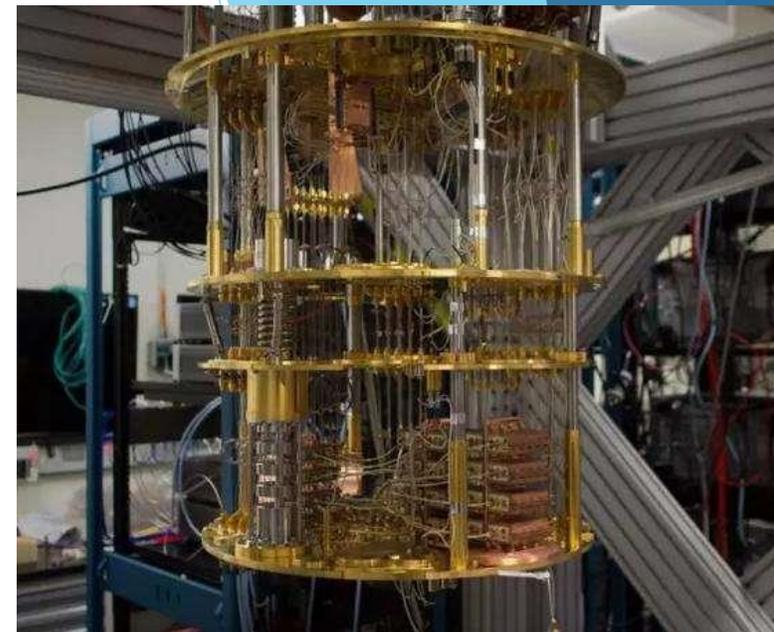
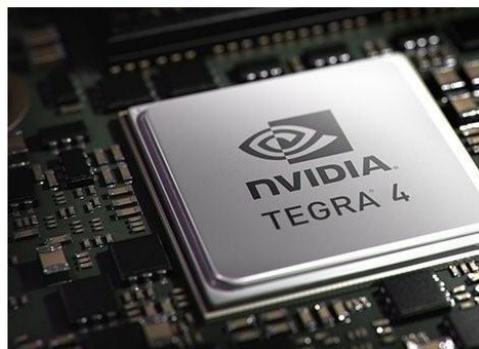
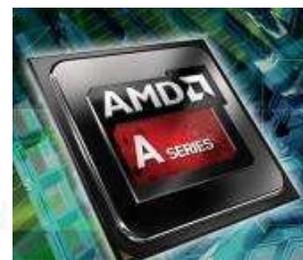




(电子管) 第一台计算机: ENICA



(硅晶体) 天河二号



量子计算发生器



墨子号, 量子纠缠

## 2016年AWS年度大会上

通过网络转移 1EB 的数据大约需要 26 年，而利用十辆 AWS 的 Snowmobile “卡车” 可以将传输时间降低到 6 个月以内。

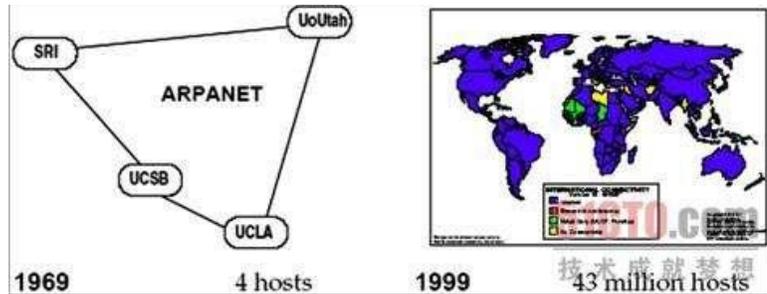


# 最早的互联网：ARPANet

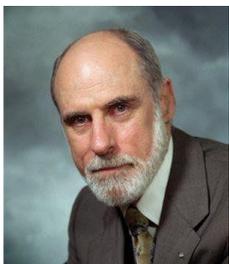
# PC互联网

# 移动互联网

上半场是门户网站的天下；



浏览器



文顿·瑟夫(Vinton G. Cerf)博士是TCP/IP协议和互联网架构的联合设计者之一，被称作“互联网之父”之一。

下半场是搜索引擎的天下，

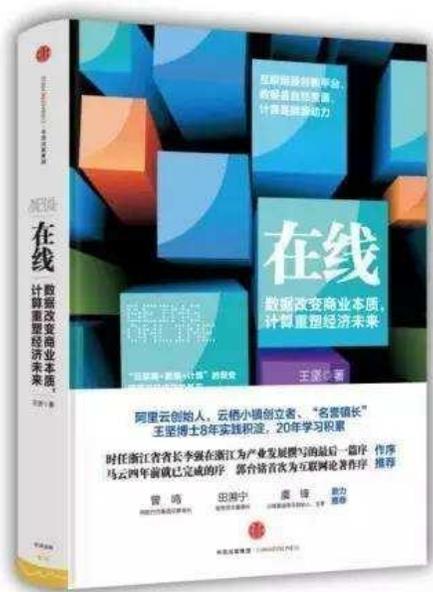
搜索引擎



蒂姆·伯纳斯·李，WWW，万维网之父

2011. 01	微信
2012. 04	朋友圈 (图文)
2012. 08	公众号 (图文)
2014. 01	微信红包
2017. 01	小程序
2017. 12	小游戏

- 从猿猴时代到猿人时代，钻木取火随时随地可以取得火光，相当于光的“在线”。
- 从农耕时代到蒸汽时代，蒸汽机可以利用蒸汽压力随时随地取得动力，是动力的“在线”。
- 从蒸汽时代到电器时代，电器可以利用电力随时随地开始工作，是能源的“在线”。
- 从互联网时代，移动设备可以随时随地使用计算和数据，是计算和数据的“在线”。



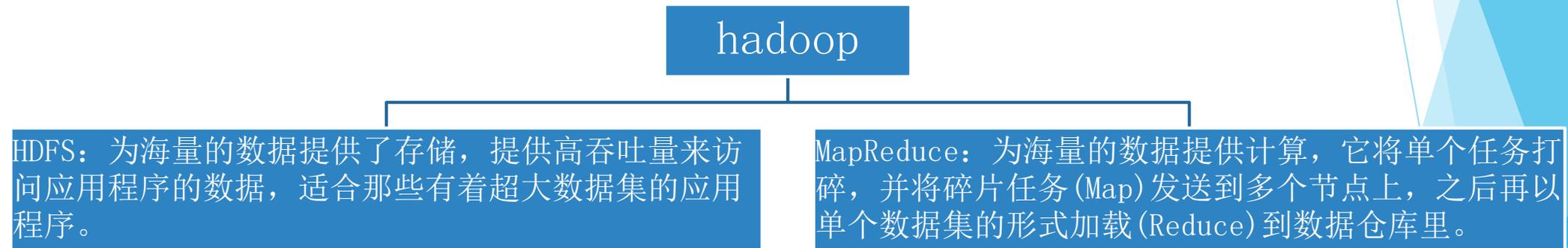
# 互联网的本质就是 数据的流动。

- 1、流动的数据消除了信息不对称。世界被紧密地对接了起来，中间环节不再被需要，“中介”成了被革命的对象，靠信息垄断赚钱的时代即将过去。电商就是典型。
- 2、流动的数据破除了时间的限制。数据的流动在线化，数据可以7\*24小时流动，而且信息是实时的。
- 3、流动的数据破除了空间的限制。没有国家和地域的限制，不再像普通门店那样只服务一个区域。
- 4、流动的数据也催生了时间的碎片化。人们在任何闲暇都可以上网，尤其是移动设备带来的便利性，这产生了“注意力经济”，哪个产品获取了更多的注意力，哪个产品就会受到青睐。

# 大数据处理的重要技术：Hadoop

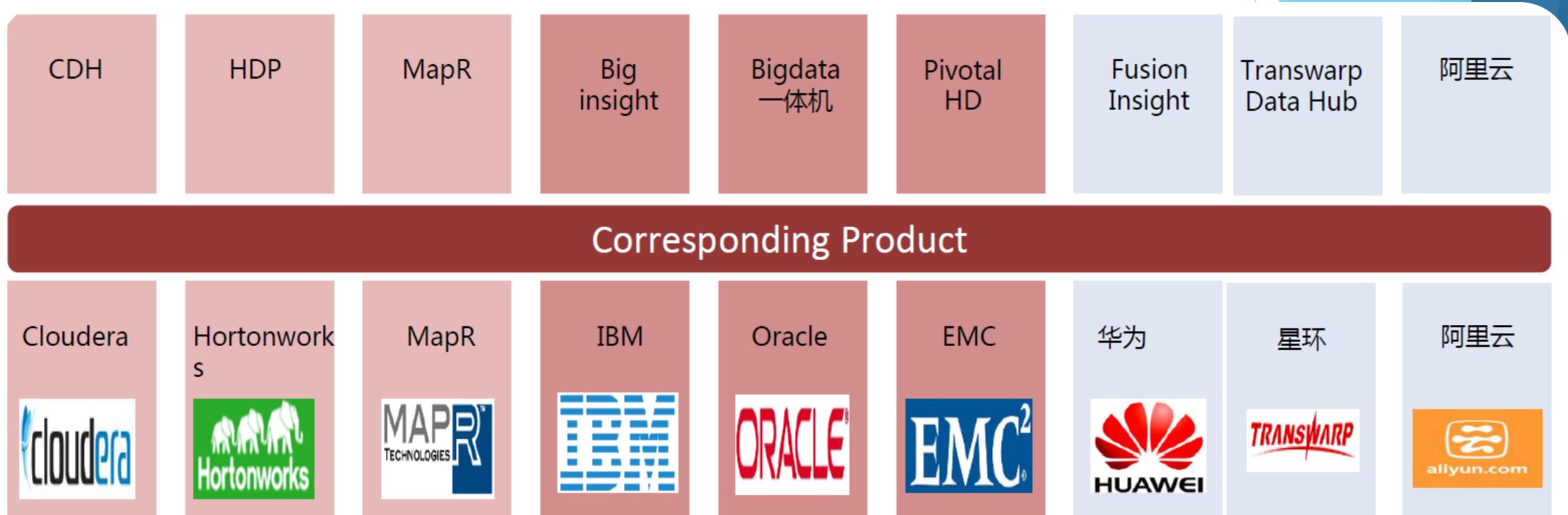
Hadoop是大数据行业应用最广的分布式系统，是一个能够对大量数据进行分布式处理的软件系统。

Hadoop框架的核心内容包括：HDFS和MapReduce。



基于以上特点，用户可以轻松地Hadoop上开发和运行处理海量数据的应用程序。

# 大数据数据存储平台能力

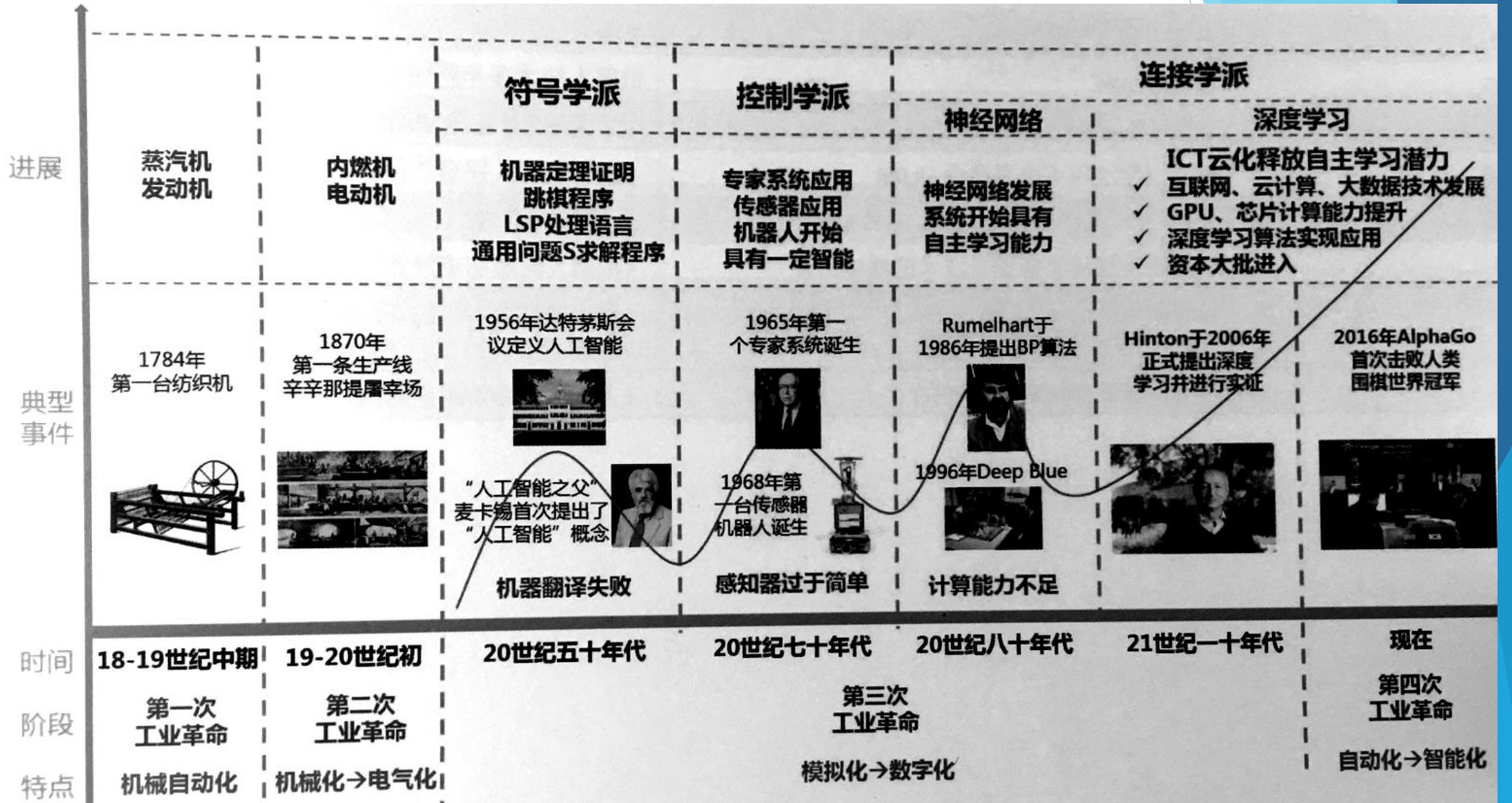


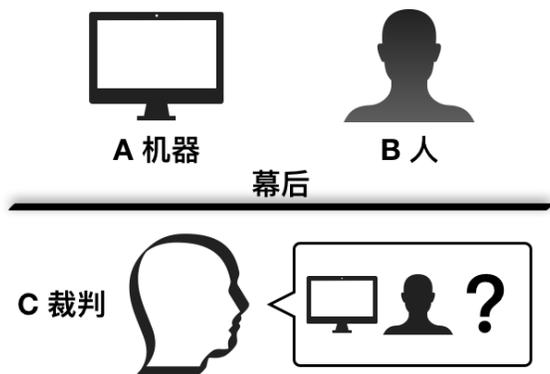
## Hadoop manufacturers

## Hadoop manufacturers



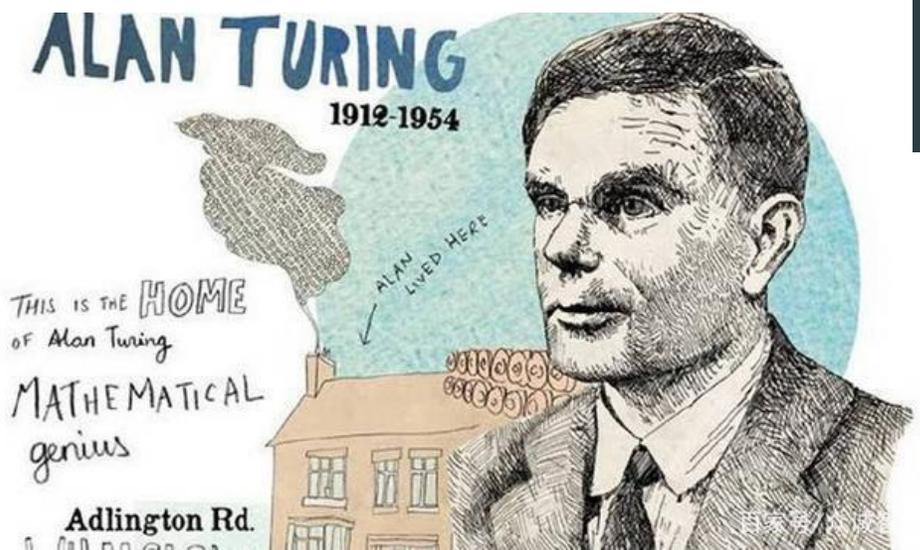
# 3、人工智能的发展



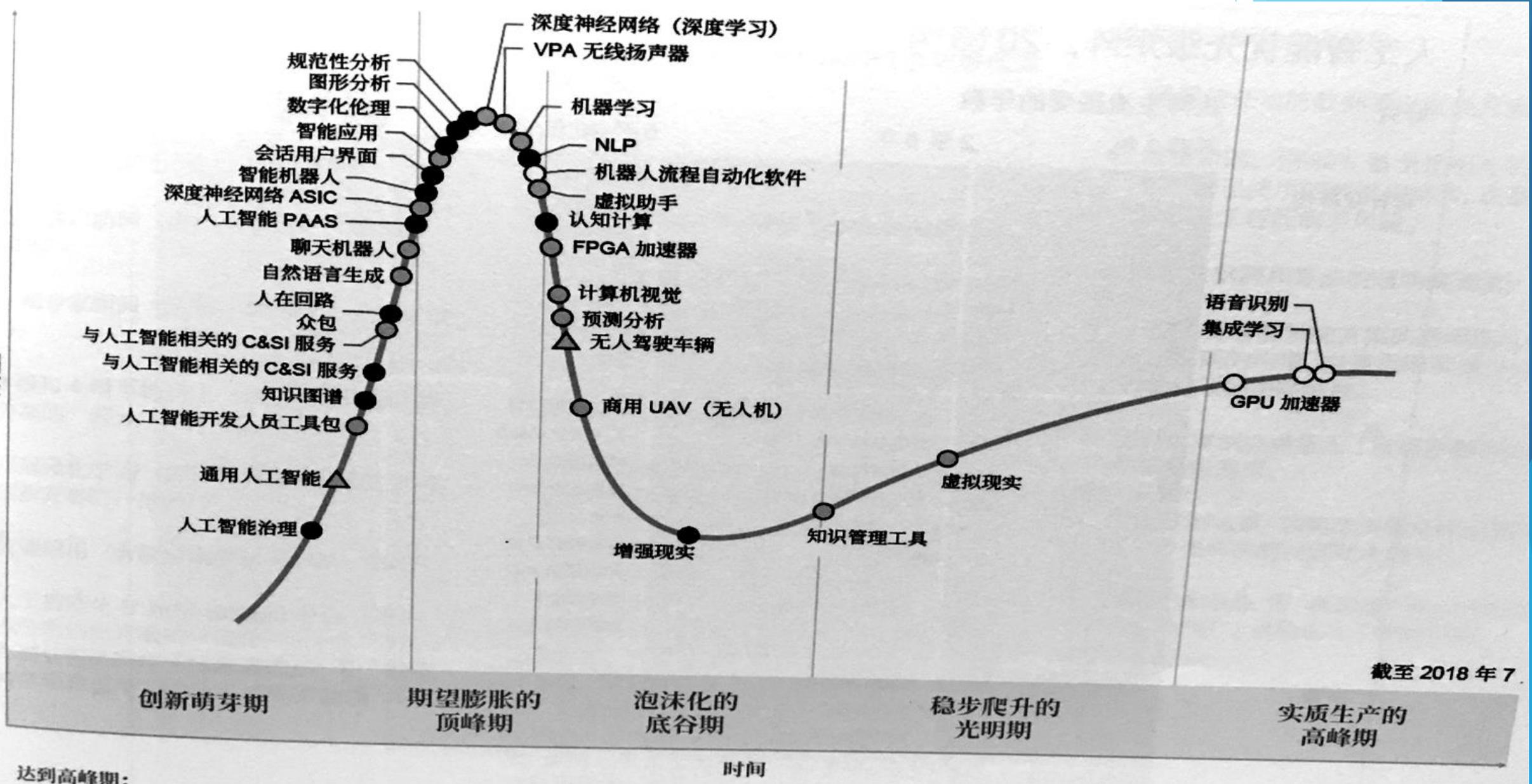


1、图灵测试（20世纪50年代）。让一台机器和一个人坐在幕后，让一个裁判同时与幕后的人和机器进行交流，如果这个裁判无法判断自己交流的对象是人还是机器，就说明这台机器有了和人同等的智能。

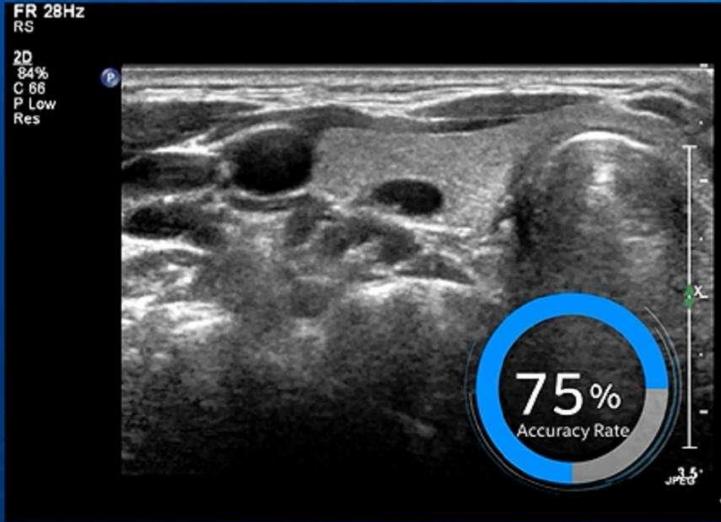
2、五个方面的机器智能（20世纪70年代）。语音识别、机器翻译、文本的自动摘要或者写作、战胜人类的国际象棋冠军、自动回答问题。



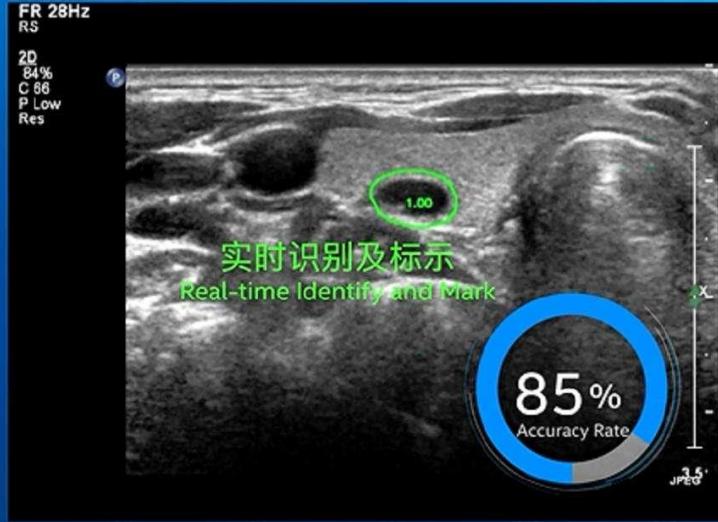
# 人工智能技术的成熟度曲线 (2018)



## 人工读片 Radiologist



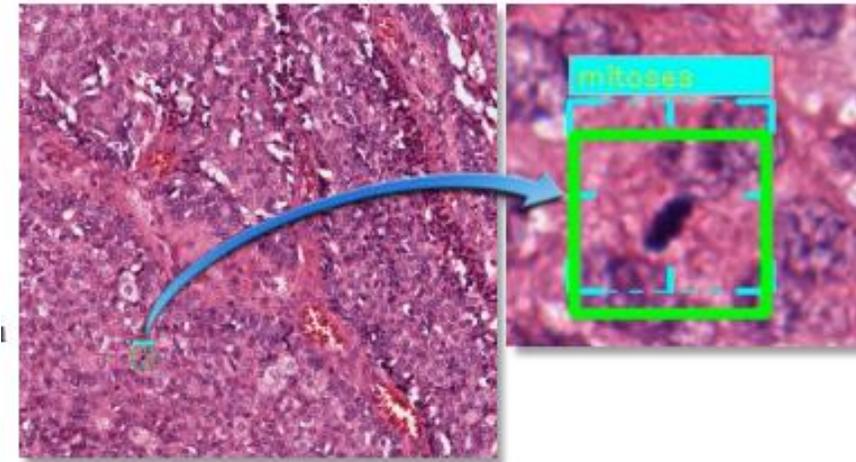
## DE超声机器人 DE Ultrasound Robot



数据来源：浙江大学数理学院研究结果  
Source: Findings from the Mathematics Department of Zhejiang University

高于人工读片的平均准确率75%  
much higher than the average of 75% of manual reading.

应用示例：乳腺组织病理切片中的癌细胞检测。绿色方框代表医生标注的癌细胞位置。



浙江大学附属第一医院针对甲状腺超声影像数据的特点对算法进行改进和优化，并利用获得的大样本对计算机进行训练，联合测试结果显示，诊断准确率可达85%以上。





儿童故事机器人



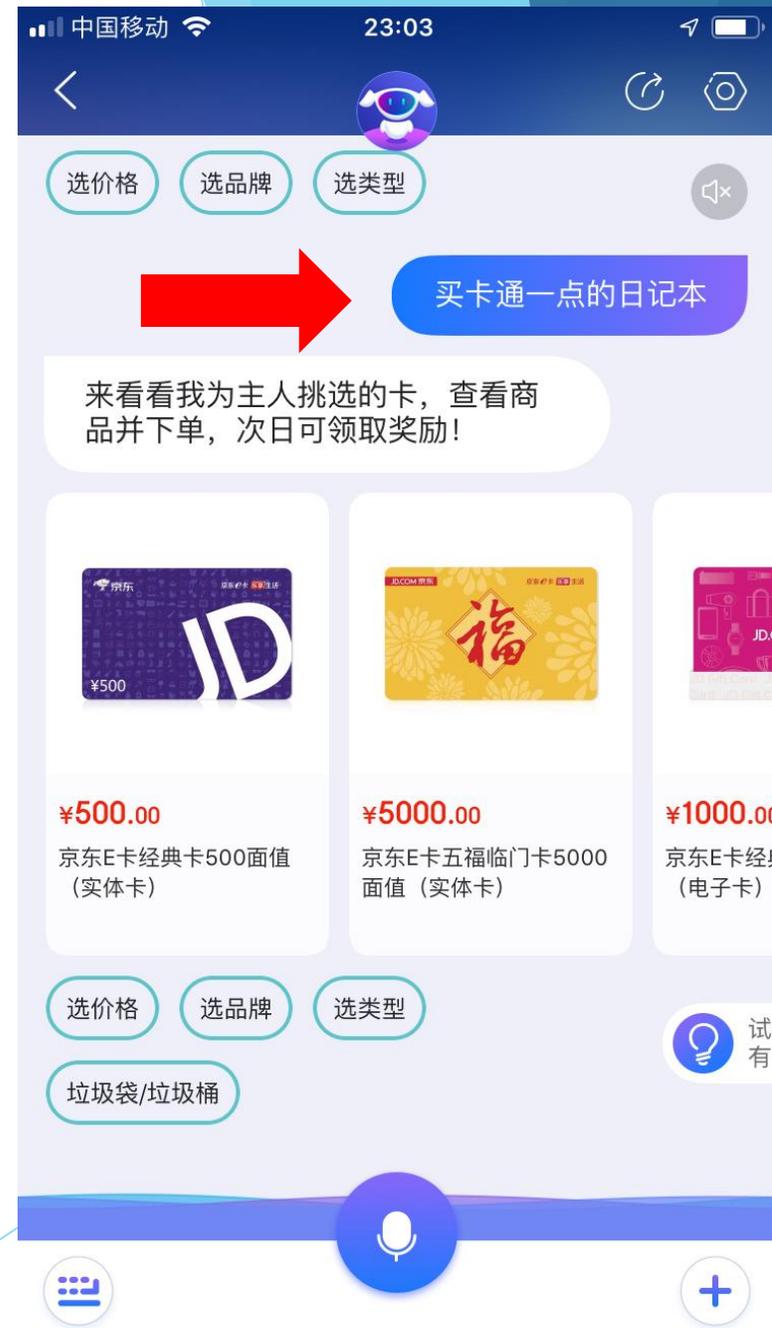
天猫精灵-购物



音箱机器人-听音乐



养成机器人-琥珀





机器人形化越来越明显。  
波士顿动力。

## 能耗是计算机发展的重要障碍。

战胜李世石的Alpha Go，拥有1202个CPU，176个GPU。按照每个CPU的功率为100W，每个GPU的功率为200W进行粗略计算，那么：

$$1202*100+176*200=173,000W=173,000J/sec$$

比赛时长以5个小时进行计算，Alpha Go的总能耗约达：

$$173,000J/sec*5hr*3,600sec/hr\approx 3,000MJ$$

而一个成年人每天所消耗的能量大约为2500kCal（大卡），换算成J（焦耳）有：

$$2500kCal*4,184J/kCal\approx 10MJ$$

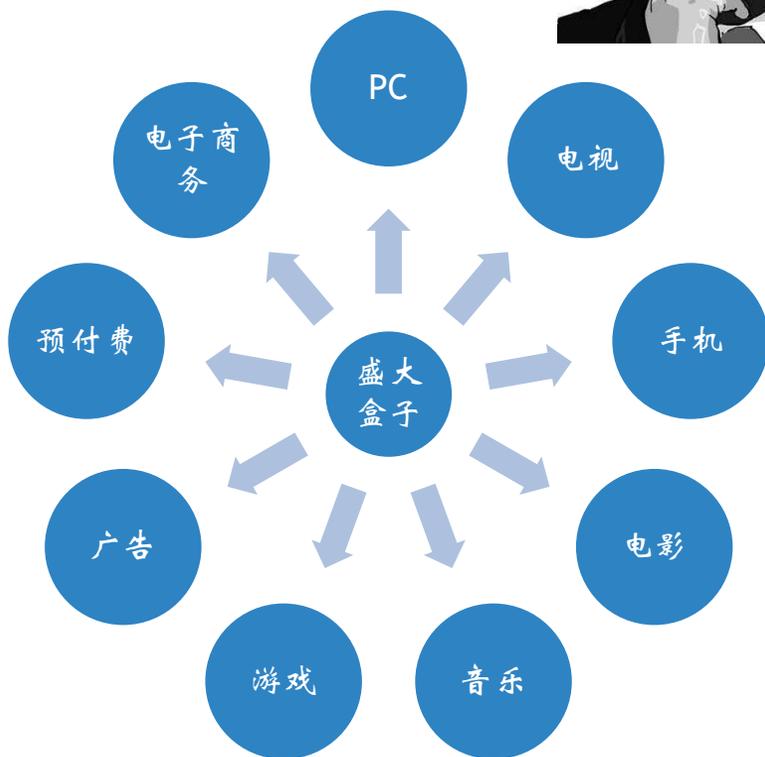
大约300倍！！

人类的大脑由1000亿个神经细胞构成，相当于银河系内的恒星数量。连接神经细胞的突触数量超过100万亿，数据存储量达到1000Tb。



# 天才敌不过天时：“半步理论”

“网络迪斯尼”计划：



2003年，他就开始搞电子支付，并在2007年将其平台化。

2004年，他就看到IP的价值，整合以“起点中文网”为核心的盛大文学。

2004年，他先于唱吧、YY直播，制作了在线K歌的游戏《巨星》。

2005年，他创全球先河，宣布旗下所有游戏免费，这种“免费玩、买道具”的模式，在当时被业界所抵制，现在却成了几乎所有网游的“致富之道”。

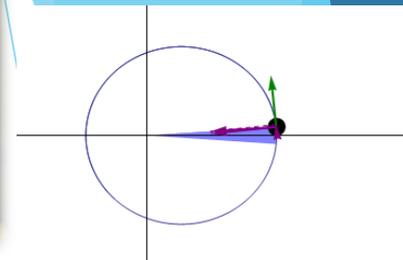
.....

他还投资了中国第一款语音IM软件TalkBox，后来的米聊、微信都是“参照”其发展起来的。

1540年哥白尼



1630年开普勒



1600布鲁诺

1630年伽利略



能力？ 时运？

天下大势、浩浩汤汤；  
顺之者昌，逆之者亡。

谁是下一个站在风口上的猪？



结论：

**数据智能**的产生是这个时代（后数字化时代）必然的结果。

## 二、大数据案例分享与应用

# 应用1、精准营销：形成以消费者为中心的产业格局

德鲁克经典问题：  
你的客户是谁？  
用户画像

实体零售遇到的问题：成本高企、竞争激烈、消费者挑剔；

## 大数据带来的机会：

全面了解消费者需求，重塑产品生产、广告营销两个环节；  
用C端需求指导F端产品生产；  
用DSP方式精准触达。



线下



线上



- 访客分析
- 回头客排行
- 实时数据
- 雷达图
- 当前在店情况
- 区域管理



身边手机排行

1	苹果	50.00%
2	中兴	50.00%

近处手机排行

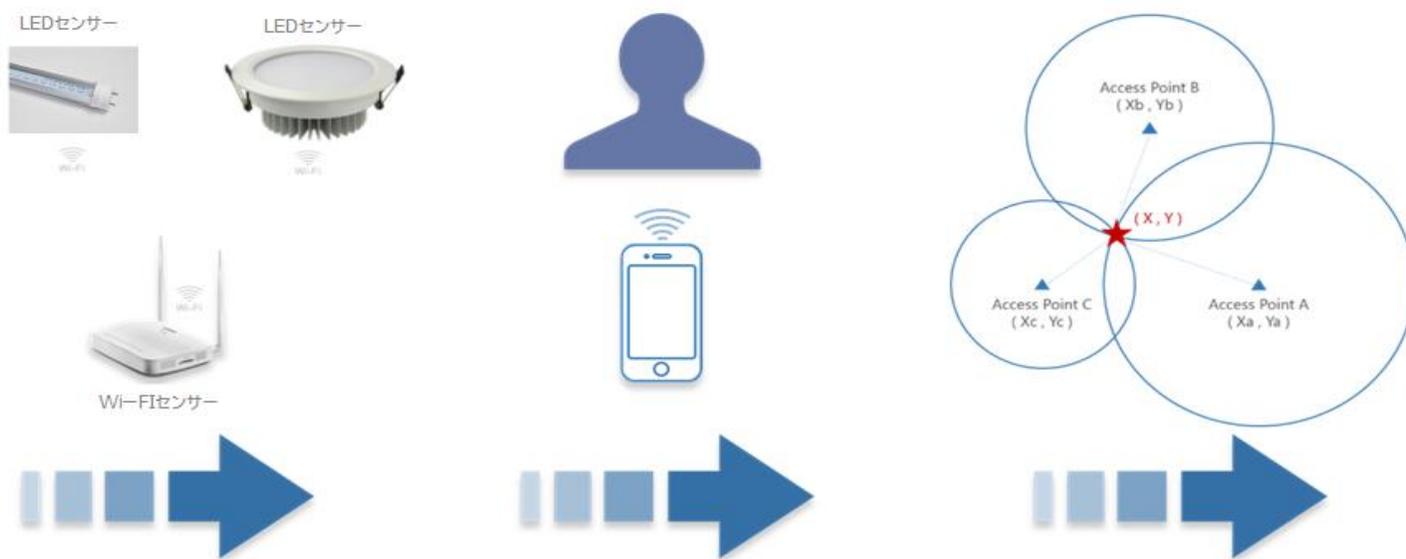
全部 身边 近处 远处  仅显示wifi注册用户

导出当前列表

序号	用户标识	手机号码	品牌	总到访次数	30天到访次数	7天到访次数	平均驻留时间
1	00*****2535	-	苹果	1	0	0	-
2	A0*****B464	-	苹果	1	0	0	-
3	58*****6399	-	索尼	1	0	0	-
4	E4*****59AD	-	三星	1	0	0	-
5	9C*****B503	-	苹果	1	0	0	-
6	28*****8C59	-	苹果	1	0	0	-
7	B4*****FC3C	-	锤子	1	0	0	-
8	98*****90D2	-	苹果	1	0	0	-
9	20*****C1E3	-	三星	1	0	0	-
10	AC*****84DD	-	苹果	1	0	0	-

# 案例：Wi-Fi围栏-人流智能监控新技术

Wi-Fi围栏基于Wi-Fi的定位算法，可以有效、精确的计算带Wi-Fi功能的智能手机（Wi-Fi需处于开启状态）的室内位置。按点布置的Wi-Fi嗅探器可以搜集周围一定范围内所有Wi-Fi信号源发出的相关信息，捕捉对象的行动踪迹，实现人流监控、剧增预测的功能。



在指定的场所内，比如商场入口，地铁站（改轨口，上下电梯口，站台等），小区入口等公共场所，布置Wi-Fi嗅探器。

顾客只要带着智能手机（只要打开Wi-Fi，不需要安装APP）。顾客所到之处通过事先设置好的Wi-Fi嗅探器来感知顾客的智能手机所发出的Wi-Fi请求连接信号，并将信号发送到云端大数据库。

在云端根据每一个时点上的环境信号动态地计算定位参数，并利用定位参数，分析计算从各个Wi-Fi嗅探器传上来的顾客智能手机的信号，以实现精确定位。（POA定位算法专利申请中）

POA  
POA室内定位

通过人流量统计功能查看指定区域，指定期限内的每一个时点的人流量热点分布图，以便实时掌握区域内大人流的剧增趋势。

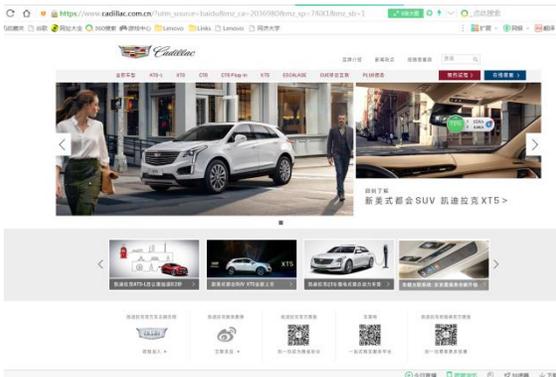


# 大数据杀熟——

“越了解你的人伤害你越深”



2018年3月31日，携程CEO孙洁等相关负责人在深圳消委会的交流会上就高价机票事件向消费者致歉。



**人口属性**  
 年龄、户籍地址、睡眠情况、工作地址、饮食习惯、活动范围、运动习惯、病史、看病时间、检查结果、诊断结果、治疗结果。

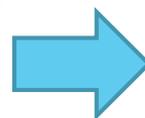
**资产情况**  
 股票、车、贵金属、房、股权、商铺、消费习惯、土地、收入。



**基本信息**  
 姓名、性别、体重、身高、婚否、职业、住址、籍贯、手机号、邮箱。

**社交图谱**  
 亲属、沟通工具。

## 网上下单数据



## 整车厂商销售团队



# 应用2、零售行业的智能化变革

人、货、场、供应链四个方面的重构



变革之由

时代与科技的发展

变革之基

智能物流的兴起与发展

变革之路

科技赋能零售产业链

# 人工智能在实体零售的应用



停车与找车

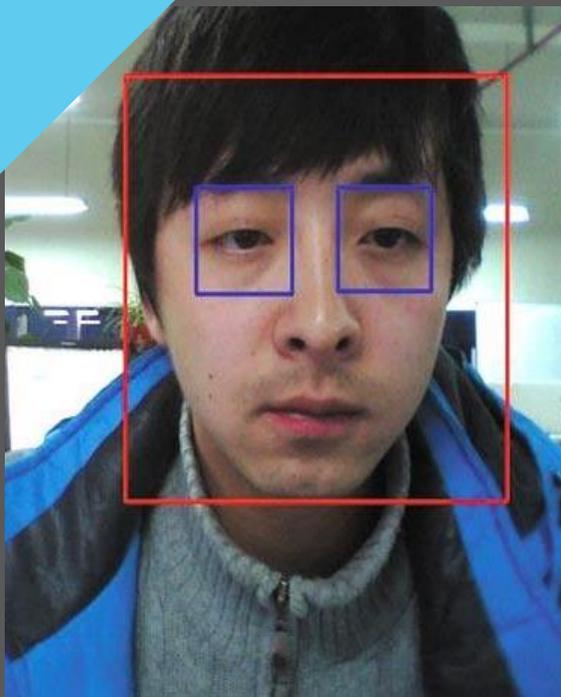


商品场内感应



顾客室内定位

# 人工智能在实体零售的应用



顾客自动识别与分析



智能购物车

比价信息

	汇源100%桃汁	
天猫		
天猫超市华南站		
【天猫超市】汇源100%纸盒桃汁1L 百分百...		12.90元
德恒北星酒类专营店		
汇源 汇源果汁200ML100% 桃汁 果蔬...		75.90元
德恒北星酒类专营店		
汇源果汁 汇源饮料 100%桃汁果汁 1L*...		140.00元
淘宝		

智能比价

# 人工智能在实体零售的应用



智能试衣间



虚拟商品墙



自助支付

# 人工智能在实体零售的应用



机器人导购



机器人库存盘点



机器人物流配送

# 区块链技术在零售业当中的使用



个人数据资源的价值变现->用户资本主义

# 应用3、城市公共安全：上海市大流量人群综合信息平台



数据服务

## 大流量人群综合信息平台

**数据采集和预处理**

手机信令数据；  
PR的WIFI数据；  
.....



**数据处理**

依托上海超算中心已有的机房，  
搭建大数据处理分析的软、硬件平台。



**基础应用**

- 1、面向各部门的基础大数据服务；
- 2、面向社会的客流大数据实时发布；
- 3、可视化呈现。



## 汇聚整合



## 高性能队列



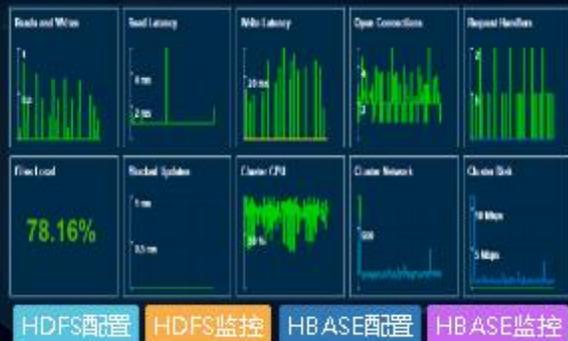
## 实时计算



## 数据共享



## 数据存储



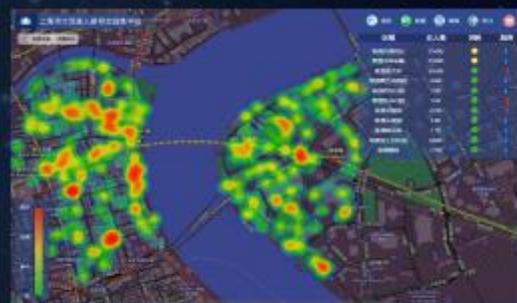
## 数据挖掘



## 可视化



## GIS地图



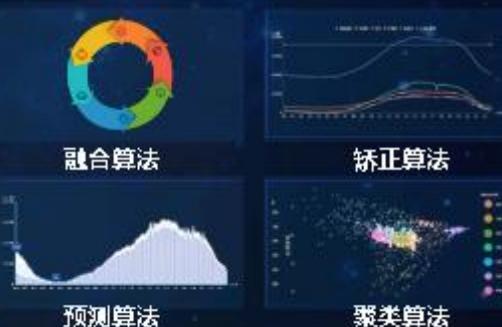
## 信息发布



## 视频监控



## 算法展现



## 系统管理

系统配置	系统管理
▶ 聚类算法配置	▶ 菜单管理
▶ 节假日配置	▶ 景区配置
▶ 修正算法配置	▶ 联通区域配置
▶ 融合权重配置	▶ 用户管理
▶ 预测算法配置	▶ 角色配置

# 模型算法研究——融合分析

- 通过多种数据源的融合之后，给出某个区域在某时间点的风险评估分数，决策者可采取不同的应对措施。比起单一数据源，多源融合风险分析打分结果更为全面、准确、合理。

重点区域客流数据融合分析	
外滩陈毅广场	76 <sup>3</sup>
外滩周边	67 <sup>1</sup>
新天地	67 <sup>1</sup>
小豫园	62 <sup>1</sup>
中豫园	62 <sup>1</sup>
小陆家嘴	61 <sup>1</sup>
外滩滨水区	41
大豫园	38
龙华寺	35
新客站	27
客运南站	18
虹桥枢纽	6
顾村公园	1



# 样本数据获取（人工实地路测）

项目组分别在2015年12月12日（周六）、2015年12月14日（周一）、2015年12月19日（周六）、2015年12月21日（周一）、2015年12月24日（周四、圣诞节）共5次分别对陈毅广场、外滩滨水区两个重点区域进行了客流的人工实地路测。

## 12月12日陈毅广场人数统计

时间	现场点人头数	照片点人头数	电子围栏数据	晚上3点基数	比例(现场)	比例(照片)
9:30	320	371	6033	2200	0.0835	0.0968
9:40	351	380	6185	2200	0.0881	0.0954
9:50	358	423	6294	2200	0.0874	0.1033
10:00	362	495	6676	2200	0.0809	0.1106
10:10	453	422	6798	2200	0.0985	0.0918
10:20	499	628	6995	2200	0.1041	0.1310
10:30	538	580	7156	2200	0.1086	0.1170
10:40	606	644	7295	2200	0.1189	0.1264
10:50	638	673	7295	2200	0.1252	0.1321
11:00	632	696	7268	2200	0.1247	0.1373

## 2015年12月12日滨水区人数统计

时间	路测人数合计	电子围栏数据	晚上3点基数	比例
9:30	1592	8252	4800	0.4612
9:40	1618	8600	4800	0.4258
9:50	1687	8964	4800	0.4051
10:00	1925	9119	4800	0.4457
10:10	2211	9471	4800	0.4733
10:20	2187	9814	4800	0.4362
10:30	2431	9986	4800	0.4688
10:40	2591	10616	4800	0.4455
10:50	2702	10694	4800	0.4584
11:00	2808	10819	4800	0.4665



# 样本数据获取（人工实地路测）

2017年6月14日、2017年8月2日、2017年10月9日分别对金牛广场和陈毅广场区域进行了客流的人工实地路测，并在路测中使用了手持移动WIFI嗅探设备。



路测方案：

路测选择半径为100m，即图中1,2,3,4,5每块区域以100m长度为准，每个人负责其中1块，最后进行汇总。路测方案中将人与车分别进行统计，小车算2人，大车算15人。

# 样本数据对比 ( 陈毅广场2017.10.09 )

2017年10月9日陈毅广场 客流监测数据源及各类模型对比

采样 时间点	实测人数 (广场 + 滨江) (不包括道路上的车辆)	电子围栏数据 (原始数据)			WiFi探测人数 (南京中山)			电子围栏活跃度算法 (已调参)		
		电子围栏人数 (原始数据)	原始数据 准确率	原始数据 偏差率	WiFi 探测人数	WiFi 准确率	WiFi 偏差率	电子围栏活跃度算法 (已调参)	活跃度算法 准确率	活跃度算法 偏差率
09:00	144	3745	2600.69%	2500.69%	144	100.00%	0.00%	168	116.67%	16.67%
09:10	163	4135	2536.81%	2436.81%	140	85.89%	14.11%	212	130.06%	30.06%
09:20	171	4369	2554.97%	2454.97%	155	90.64%	9.36%	218	127.49%	27.49%
09:30	224	4594	2050.89%	1950.89%	144	64.29%	35.71%	238	106.25%	6.25%
09:40	229	4850	2117.90%	2017.90%	160	69.87%	30.13%	258	112.66%	12.66%
09:50	254	5032	1981.10%	1881.10%	189	74.41%	25.59%	256	100.79%	0.79%
10:00	255	5181	2031.76%	1931.76%	180	70.59%	29.41%	277	108.63%	8.63%
10:10	310	5328	1718.71%	1618.71%	188	60.65%	39.35%	289	93.23%	6.77%
10:20	312	5387	1726.60%	1626.60%	207	66.35%	33.65%	296	94.87%	5.13%
10:30	358	5506	1537.99%	1437.99%	200	55.87%	44.13%	294	82.12%	17.88%
10:40	392	5694	1452.55%	1352.55%	199	50.77%	49.23%	325	82.91%	17.09%
10:50	355	5659	1594.08%	1494.08%	217	61.13%	38.87%	320	90.14%	9.86%
11:00	359	5662	1577.16%	1477.16%	203	56.55%	43.45%	323	89.97%	10.03%
11:10	411	5620	1367.40%	1267.40%	204	49.64%	50.36%	333	81.02%	18.98%
11:20	358	5703	1593.02%	1493.02%	216	60.34%	39.66%	342	95.53%	4.47%
11:30	360	5632	1564.44%	1464.44%	241	66.94%	33.06%	337	93.61%	6.39%
11:40	362	5706	1576.24%	1476.24%	241	66.57%	33.43%	355	98.07%	1.93%
11:50	337	5710	1694.36%	1594.36%	200	59.35%	40.65%	345	102.37%	2.37%
12:00	395	5698	1442.53%	1342.53%	185	46.84%	53.16%	354	89.62%	10.38%
<b>均值</b>	<b>303</b>	<b>5222</b>		<b>1727.00%</b>	<b>190</b>		<b>34.00%</b>	<b>292</b>		<b>11.00%</b>

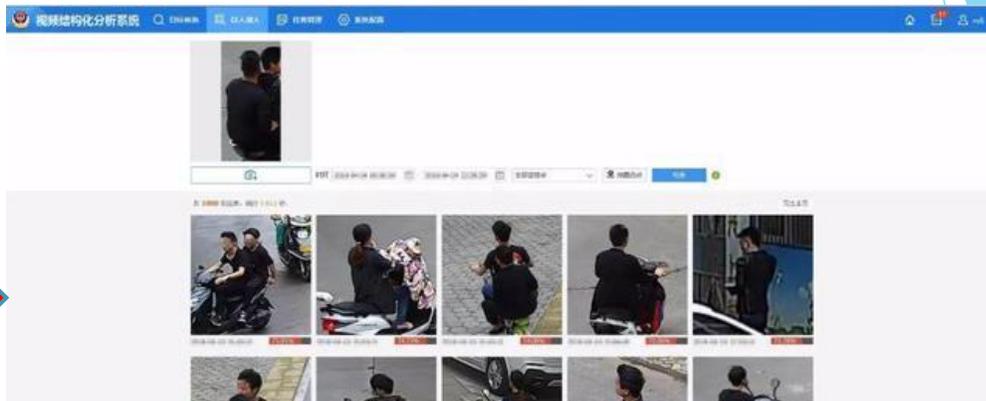
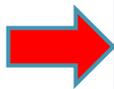
# 应用4、城市综合治理：静安区“151”工程

静安区“151项目”被国家发改委列为2018年数字经济试点重大工程支持项目，其主要内容是“1+5+1”：

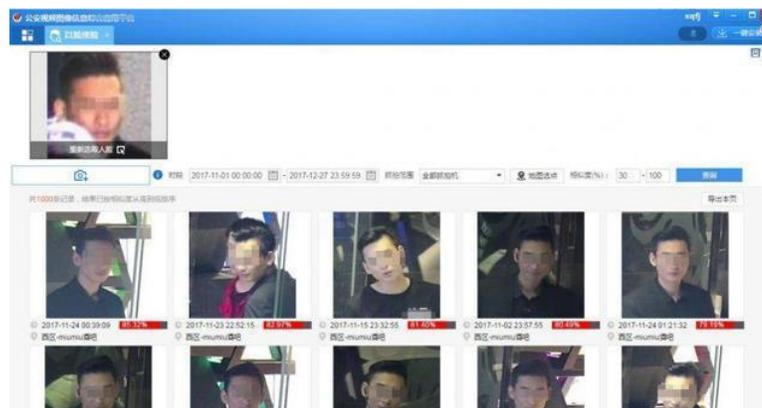
- 1、前一个“1”是建设一张新型城域物联感知网络；
- 2、“5”是完成在交通、健康医疗、食品安全、环保、城市公共设施等五大民生热点领域20万套感知设备的部署，构建数据自动采集、充分共享的城市精细化管理“神经元末梢”；
- 3、后一个“1”是以“大数据+”为创新手段，升级现有网格中心平台，建成1个具有城市数据综合运营能力、大数据分析挖掘能力和预测预警能力的城市智能综合运营管理中心，即“城市智慧大脑”。



# 智能破案：从一个背影抓捕犯罪嫌疑人



9小时破案!



比如，通过物联感知设备对辖区内井盖、电箱、垃圾桶、道路扬尘等公共部件进行实时监控及报警；通过视频流分析事件智能感知，收集路面人流数据、路段的车流情况，车速、车型、违章等数据，对辖区的人、车做到全面的管理和分析；通过网格城运、市政市容作业、城管作业、景观灯光管理、静态停车、重点场所消防监控、环保气象、商业数据等多维城市管理数据的汇聚和叠加，以及与BIM建筑物内部空间模型等城市管理地图的融合，实现跨区域快速研判、精准指挥、高效处置。



# 应用5、医疗健康：构建个性化医疗模式



从挂号网 —————>微医网

2014年10月，1亿美金（来自腾讯）



青苹果



丁香医生  
WWW.DXY.COM



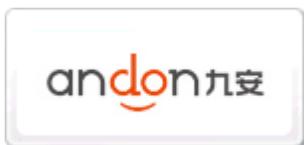
好大夫在线  
WWW.HAODF.COM



名医在线

医生资源

病人



→ 形成大数据 → 个性化医疗

北京市卫计委下发通知，3月25日前，明令严禁医生与商业公司合作  
挂号加号。



中国药品电子监管码积累的大量数据，生产、存储、流通等



百度医疗产品“百度医生”接入的公立医院资源占全国的三分之一



百度医生吧

2015年全球精准医疗市场规模近600亿美元。据机构预测，未来五年，全球精准医疗市场的规模将继续快速增长，2015-2020年期间增速可达15%，是医药行业增速的3-4倍。

上游是测序仪器提供商，中游是测序服务提供商，下游是测序服务的应用市场。

地区	测序服务机构												
北京	诺禾致源	安诺优达	贝瑞和康	圣谷同创	博奥生物	百迈克生物	康普森	嘉宝仁和	源宜基因	博淼生物	圣庭集团	中美泰和	斯科尔基因
	华牛生物	微旋基因	基云惠康	爱普益	迈基诺基因	量化健康	诺赛基因	毅新兴业	博恒生物	百麦华康	华生恒业	路思达	鑫诺美迪
	中科紫馨	海克维尔	瑞德百奥	英木和	溯源精微	华奕生物	奥维森	布斯坦	信诺佰世	银河基因	华诺时代	泛生子	
上海	药明康德	云健康	派森诺生物	晶能生物	美吉生物	宝腾生物	凡迪生物	佰真生物	南方基因	烈冰科技	生工生物	鼎晶圣物	锐羿基因
	欧易生物	翰宇生物	泛亚基因	尤妮曼	联合基因	吉玛生物	康成生物	赛安生物	吉凯基因	上海敏芯	阿趣生物	博苑生物	丰核信息
	生咨生物	英拜生物	凌科生物	泉麦生物	卓立生物	达迈生物	基因科技	基龙生物	源奇生物	赛优生物	希匹吉生物	百世嘉	派航生物
	伯豪生物	祥音生物	锐羿生物	惠研生物	嘉因生物	允英医疗	虹舜生物	捷易生物	思路迪-埃提斯		illumina中国		TAAG Genetics
深圳	华大基因	千年基因	博大维康	易基因	海普洛斯	裕策生物	蓝图基因	普元科技	早知道科技	英马诺生物	恒创基因	锐澳康晨	华因康基因
广州	达安基因	托普基因	锐博生物	燃石科技	基迪奥生物	永诺生物	坤图生物	英格生物	金域检验	瑞科基因	赛哲生物	洪祥生物	
杭州	贝达药业	谷禾生物	浙江天科	中翰金诺	杭州英睿	壹基金	然纳生物	晶佰生物	美格生物	联川生物	奥拓生物	艾迪康医学	迪安诊断
武汉	菲沙基因	康圣环球	贝纳基因	生命之美	数桥科技	锦奥生物	大众源生						
苏州	帕诺米克生物	金唯智生物	贝斯派生物	天昊生物	赛业生物	Qiagen	苏州生物医药创新中心						
其它	中宜金大(江苏宜兴)		亿康基因(江苏泰州)		所罗门兄弟医学(江苏盐城)		世和基因(江苏南京)		迪康金诺(江苏南京)		广而生物(江苏南京)		
	苏博生物(江苏宿迁)		锐创生物(江苏绍兴)		健海生物(河北石家庄)		盘古基因(天津)		天津生物芯片(天津)		国信凯尔(山西太原)		
	先导药物(四川成都)		湘雅医学检验所(湖南长沙)		博川基因(湖南长沙)								

# 应用6、金融领域

## (1) 大数据与支付



2016年4月26日, 蚂蚁金服B轮融资后, 估值3895亿人民币。比肩Uber (625亿美元), 交通银行A股市值4300亿人民币。



# 支付一切



# 人类进入无现金社会

丹麦政府自2016年起便已开始基本进入“无现金社会”。这个国家总人口为560万人，其中，有200万人使用移动支付服务，因此，丹麦政府从2016年开始实施“无纸币政策”。

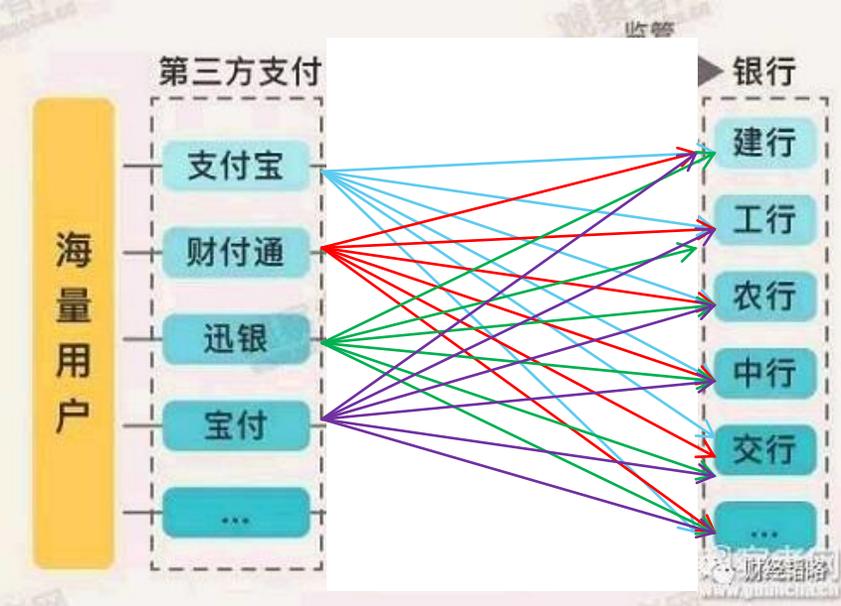


根据斯德哥尔摩KTH皇家理工学院技术人员观察，瑞典有望成为世界上第一个无现金社会。根据瑞典央行的数据，非现金交易平稳增长，比过去十年多出10%，达到了92%的高峰。

莫迪在推特（twitter）发文说，转向数字化交易，可以终结腐败。希望“DigihanMela”活动提高印度民众对数字支付的认识，帮助人们开设账户、登记取得生物辨识卡“Aadhaarcad”，使用数字交易。

**2017年8月4日，央行宣布：**  
**从2018年6月30日起，类似支付宝、财付通等第三方支付公司受理的，涉及银行账户的网络支付业务，都必须通过“网联支付平台”处理。**  
**截止10月15日，已有20余家支付平台接入网联系统。**

序号	股东名称	本次实缴出资额(万元)	股权比例	背景
1	中国人民银行清算总中心	12000.00	12.00%	央行直属
2	梧桐树投资平台有限责任公司	10000.00	10.00%	国家外汇管理局
3	财付通支付科技有限公司	9610.00	9.61%	腾讯
4	支付宝(中国)网络技术有限公司	9610.00	9.61%	阿里巴巴
5	网银在线(北京)科技有限公司	4710.00	4.71%	京东
6	上海黄金交易所	3000.00	3.00%	其他国家机构
7	银行间市场清算股份有限公司	3000.00	3.00%	其他国家机构
8	中国银行间市场交易协会	3000.00	3.00%	其他国家机构
9	中国印钞造币总公司	3000.00	3.00%	央行直属
10	中国支付清算协会	3000.00	3.00%	央行直属
11	天翼电子商务有限公司	2770.00	2.77%	中国电信
12	快钱支付清算信息有限公司	2450.00	2.45%	万达
13	平安付科技服务有限公司	2450.00	2.45%	平安集团
14	北京百付宝科技有限公司	2420.00	2.42%	百度
15	联通优势电子商务有限公司	1990.00	1.99%	海立美达
16	中移电子商务有限公司	1640.00	1.64%	中国移动
17	银联商务有限公司	1660.00	1.55%	银联
18	通联支付网联服务股份有限公司	1280.00	1.28%	
19	易宝支付有限公司	1210.00	1.21%	
20	顺丰恒通支付有限公司	1200.00	1.20%	顺丰
21	上海汇付数据服务有限公司	1180.00	1.18%	
22	网易宝有限公司	1100.00	1.10%	网易
23	上海盛付通电子支付服务有限公司	1080.00	1.08%	盛大
24	易智付科技(北京)有限公司	1060.00	1.06%	
25	深圳市美的支付科技有限公司	970.00	0.97%	美的
26	银盛支付服务股份有限公司	910.00	0.91%	
27	拉卡拉支付股份有限公司	900.00	0.90%	
28	深圳瑞银信息技术有限公司	900.00	0.90%	
29	迅付信息科技有限公司	860.00	0.86%	
30	连连银通电子支付有限公司	840.00	0.84%	
31	联通支付有限公司	840.00	0.84%	
32	上海瀚银信息技术有限公司	840.00	0.84%	
33	宝付网络科技(上海)有限公司	830.00	0.83%	漫道金服
34	汇元银通(北京)在线支付技术有限公司	830.00	0.83%	
35	杉德支付网络服务发展有限公司	810.00	0.81%	
36	智付电子支付有限公司	770.00	0.77%	
37	海南新生信息技术有限公司	750.00	0.75%	
38	深圳市快付通金融科技服务有限公司	750.00	0.75%	
39	汇潮支付有限公司	710.00	0.71%	
40	浙江唯品会支付服务有限公司	690.00	0.69%	
41	东方电子支付有限公司	680.00	0.68%	
42	上海付费通信息服务有限公司	640.00	0.64%	
43	国付宝信息科技有限公司	620.00	0.62%	
44	易联支付有限公司	410.00	0.41%	
45	捷付睿通股份有限公司	140.00	0.14%	小米
	合计	100000.00	100.00%	



**中国人民银行 (支付结算司)**

银支付〔2017〕209号

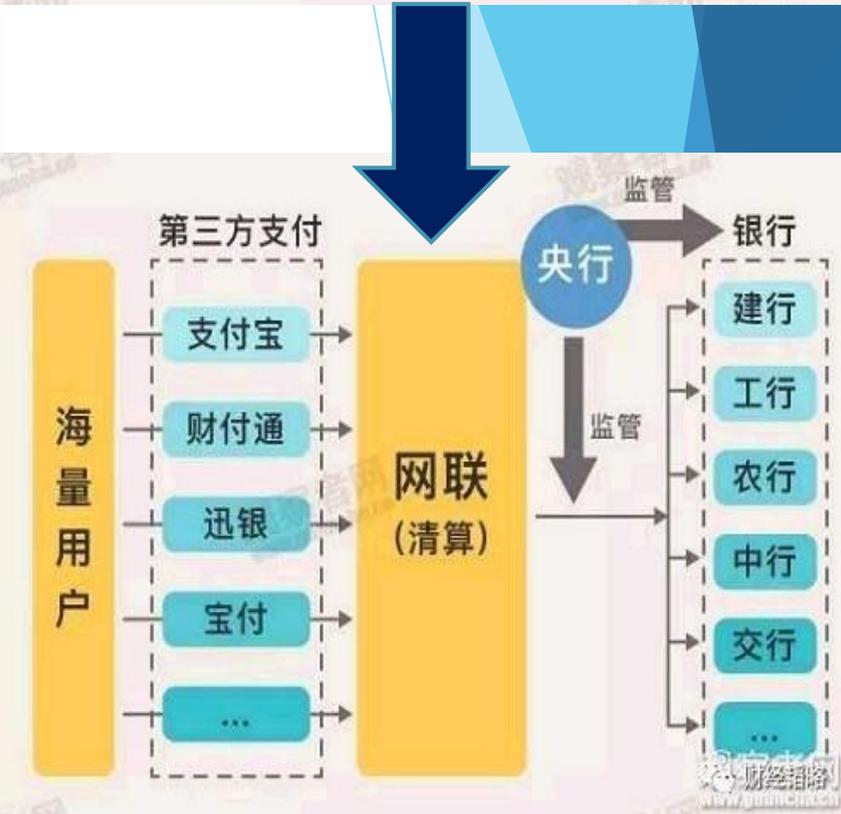
**中国人民银行支付结算司关于将非银行支付机构网络支付业务由直连模式迁移至网联平台处理的通知**

中国人民银行上海总部金融服务一部，各分行、营业管理部、省会(首府)城市中心支行，各副省级城市中心支行支付结算处(会计财务处)；各国有商业银行、股份制商业银行，中国邮政储蓄银行运营管理部；各非银行支付机构：

根据党中央、国务院关于互联网金融风险专项整治的工作部署，人民银行指导支付清算协会建设“非银行支付机构网络支付

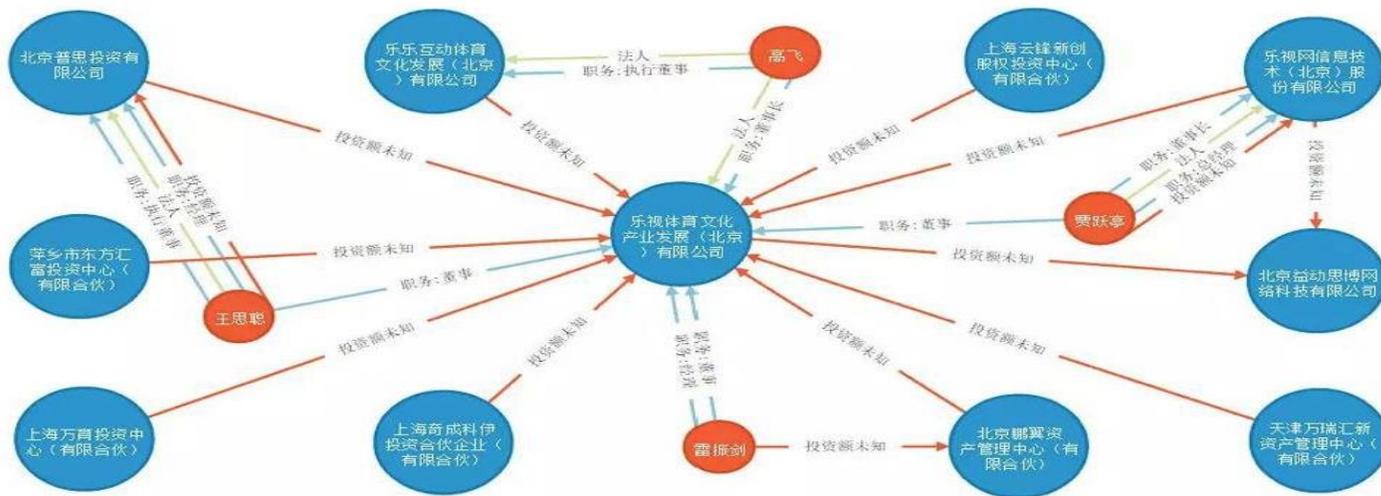
- 1 -

财经丝路



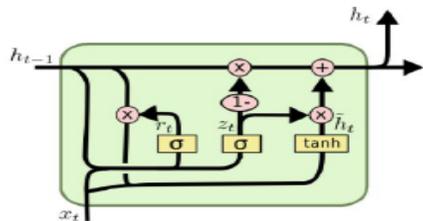
## (2) 大数据与银行风控

# 企业征信信息查询服务平台

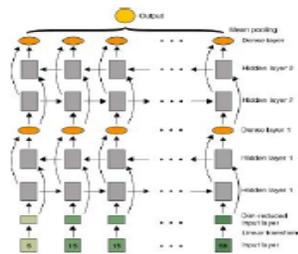


# 企业风险评估大数据框架图

## 企业经营趋势(风险评估)



## 语义搜索



## 企业知识图谱



语义挖掘

语义解析

数据源

细分行业分析

所属热点话题

唯一实体识别

语义地址解析

上下游关系图谱

任职关系图谱

财报语义解析

情感倾向

语义事件解析

担保关系图谱

资金流向图谱

新闻

贴吧

法院

环保

工商

财报

招投标

...

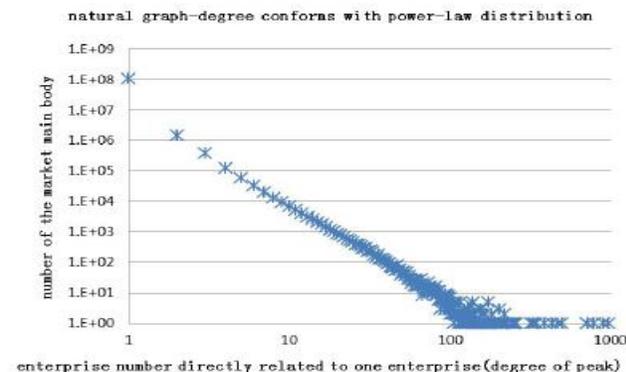
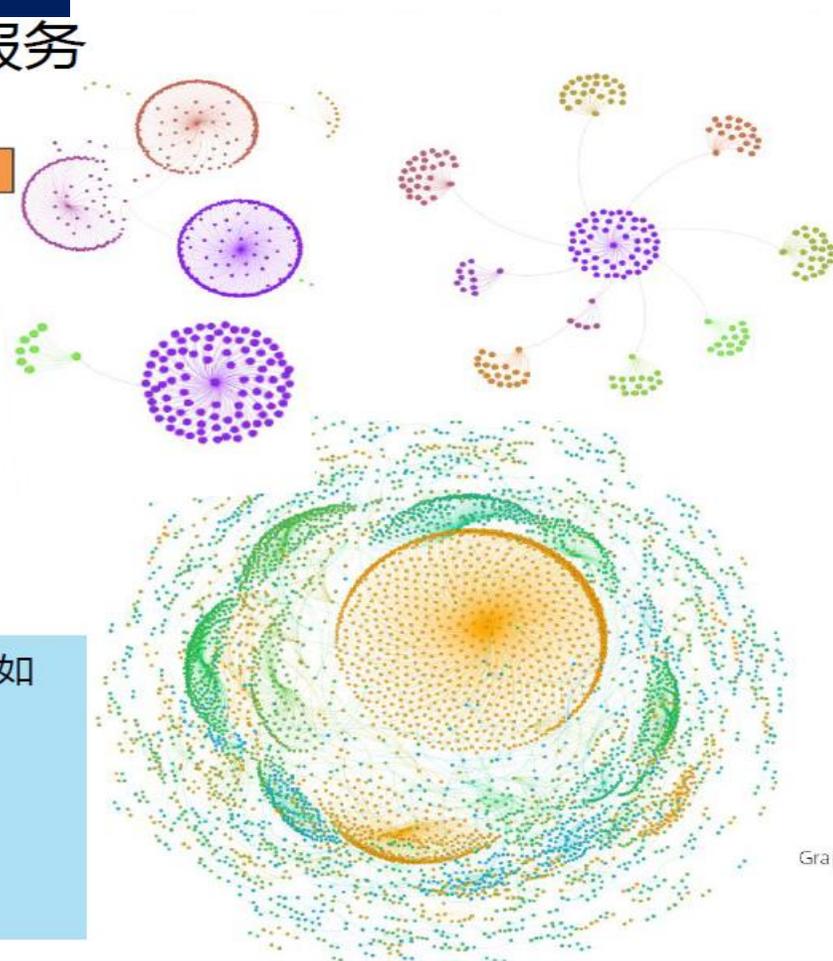
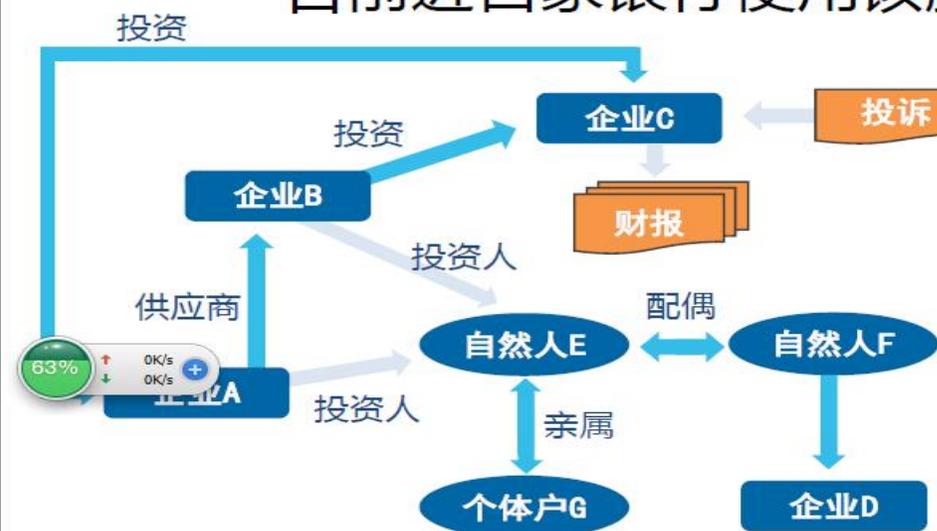
基础信息

账务流水

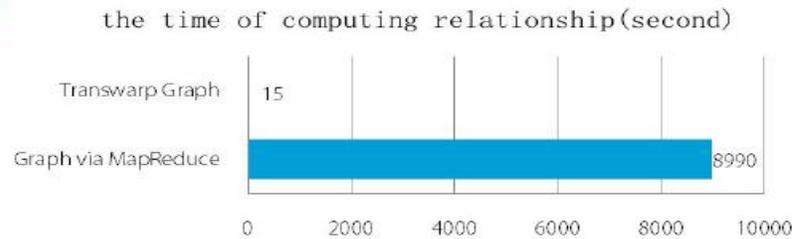
贷款

# 企业投资任职关系分析

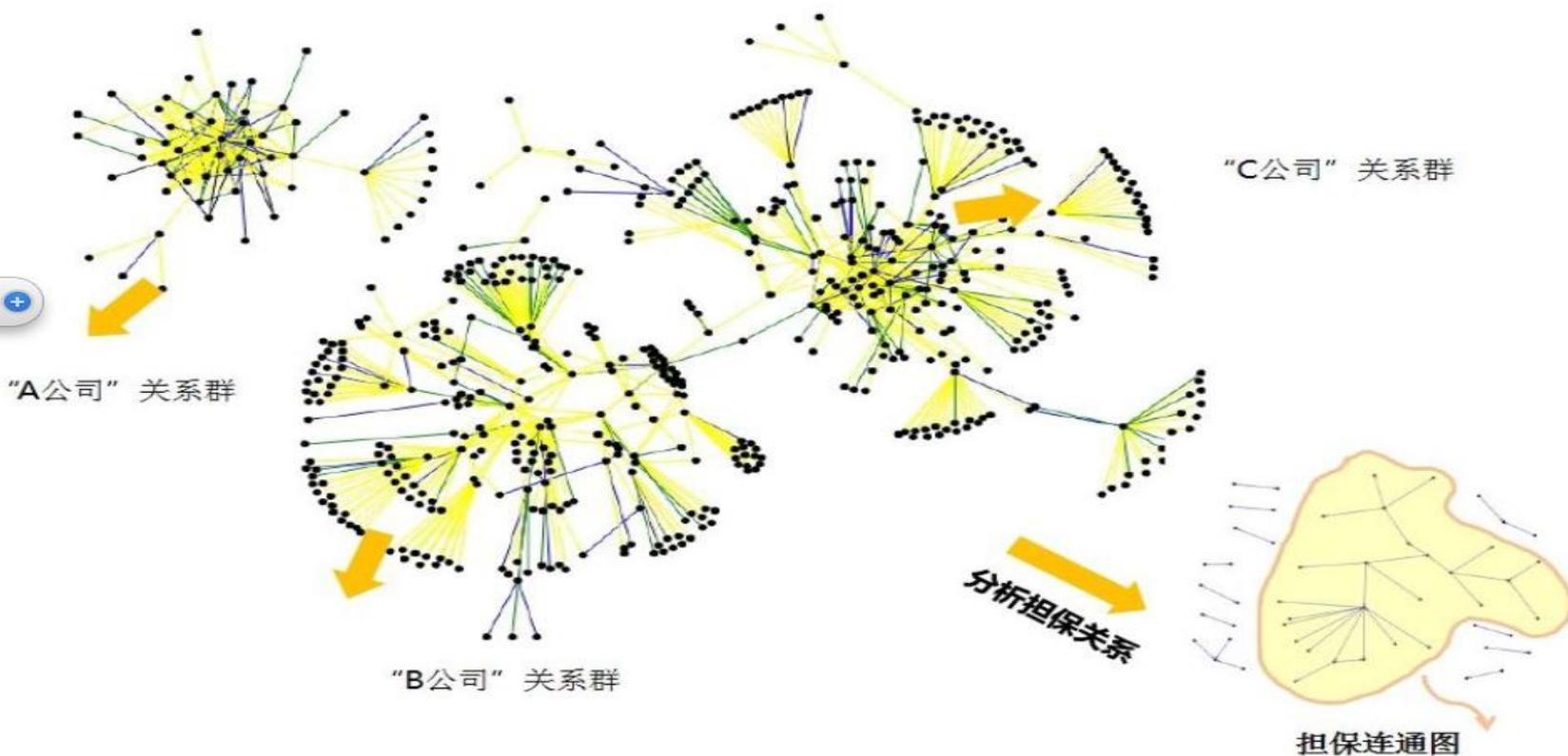
目前近百家银行使用该服务



- ((1)) 多家企业之间是否有关联？如何关联？关联关系如何变化？
- ((2)) 供应链管理？如何变化？
- ((3)) 企业信用分析？



# 某银行客户关联关系的分析



- 深度挖掘企业与关联企业、企业与关联个人、个人与关联个人之间的关系。
- 将关联关系网络形成直观的关系图
- 用于市场营销、潜在客户开拓、风险预警等

### (3) 大数据与个人征信

#### 用户授权数据



#### 聚信立报告

通话情况分析、联系人分析、出行分析、消费习惯和消费能力分析



#### 学历学籍

学历数据、学籍数据、学生身份验证



#### 信用卡账单

支持中国主要银行信用卡的额度和账单分析，支持各大主流邮箱



#### 公积金社保

支持全国三百多个城市的公积金和社保缴费数据，工作和房贷分析



#### 保险网站

支持中国主流的保险网站，车险财险数据



#### 征信网站

支持个人征信报告的查询



#### 简历网站

支持中国主流的简历网站，个人技能和工作分析



各种微小贷



**2015年1月5日**，央行允许**8家**社会机构开展个人征信业务。芝麻信用管理有限公司、腾讯征信有限公司、深圳前海征信中心股份有限公司、鹏元征信有限公司、中诚信征信有限公司、考拉征信有限公司、中智诚征信有限公司、北京华道征信有限公司。

经过两年的试验，没有一家合格。原因是：

第一、数据不共享。每一家都想追求依托互联网形成自己的业务闭环。

第二、不具有第三方征信独立性。这**8家**机构各自依托某一个企业或者企业集团发起创建。

第三，信用评分不规范。这**8家**机构对征信的基本理念和基本规则了解不够，而且也不太遵守，根据各自掌握的有限的信息进行不同形式的信用评分并对外进行使用，存在信息误采误用问题。

2018年2月22日，央行发布信息，百行征信有限公司个人征信业务获得央行许可，个人征信业务服务对象为从事互联网金融个人借贷业务的机构，另外还包括传统金融机构、金融监管相关部门及个人信息主体，牌照有效期至2021年1月31日。

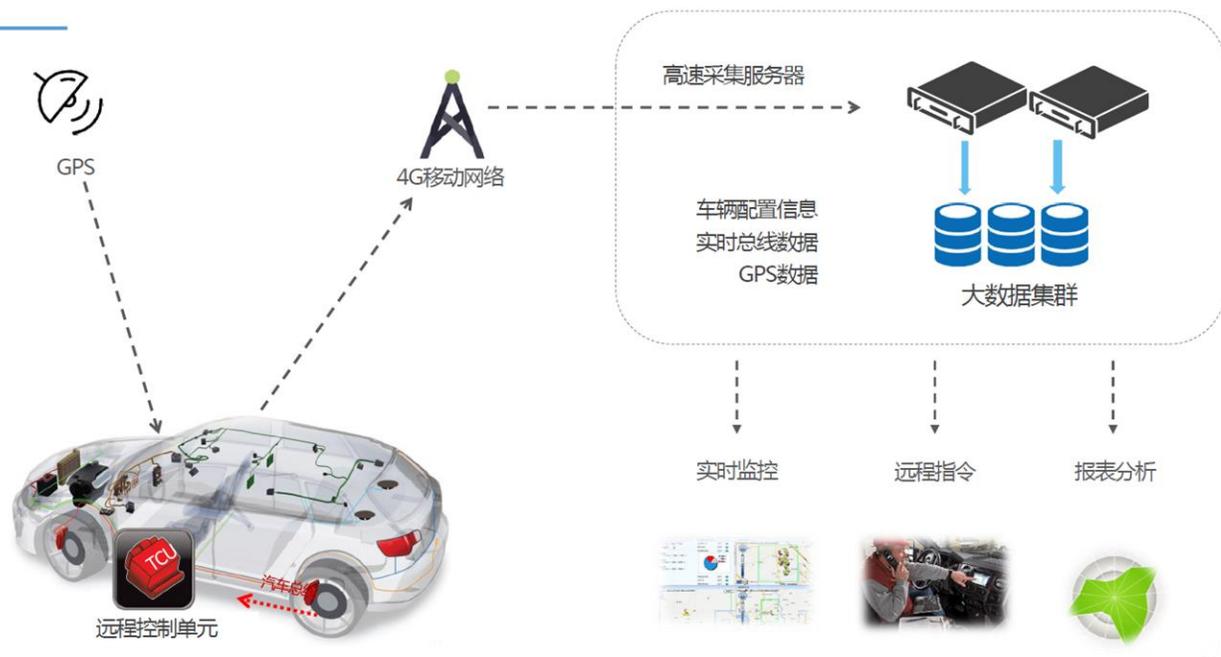
## 个人征信 牌照



名称	股份
中国互联网金融协会	36%
芝麻信用管理有限公司	8%
腾讯征信有限公司	8%
深圳前海征信中心股份有限公司	8%
鹏元征信有限公司	8%
中诚信征信有限公司	8%
考拉征信有限公司	8%
中智诚征信有限公司	8%
北京华道征信有限公司	8%

## (4) 大数据与保险

车辆远程管理系统



车辆保险：按日结算

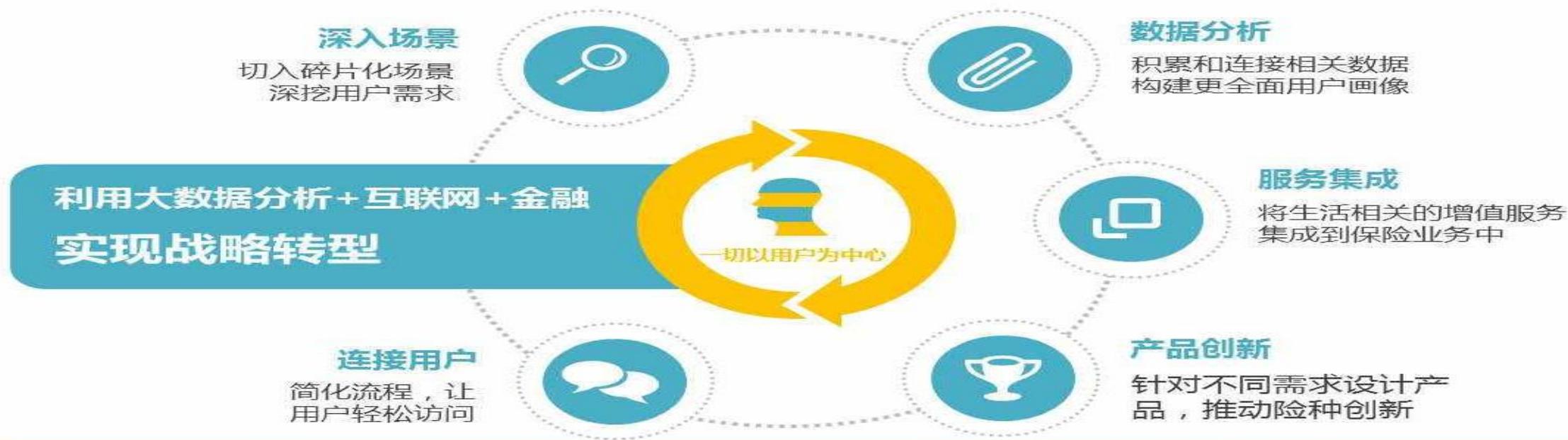


GPS定位

UBI (User Behavior Insurance) 基于用户驾驶习惯的保险。  
现在美国保险公司甚至通过和车企合作, 得到车主驾车习惯的数据, 那些喜欢大脚踩油门的车主恐怕会面临保费上浮的可能



## UBI能帮助保险公司探索创新业务模式



在产品、服务、商业模式、体制机制的改革创新将提升保险公司的核心竞争力。



2017年11月，微信“九宫格”凑满，首先在微信“保险服务”中推出的是“微医保·住院医疗险”。

目前重疾险已上线，未来车险、驾乘意外险、航意航延险等都将陆续推出来。

凡是“微保”投保用户，在微信运动上每天步行8000步就可以领取1元的运动鼓励金，保障期内最高可领取100元。深圳投保用户累计行走54亿步，可绕地球67圈，已累计领走21万元运动鼓励金。

## 重要告知

1. 此产品由中国人民财产保险股份有限公司倾情提供并承保。
2. 此产品领取成功后向用户发送电子保单号电子保单。
3. 保单形式：网上投保为您提供电子保单，根据《中华人民共和国合同法》第十一条规定，数据电文是合法的合同表现形式，电子保单与纸质保单具有同等法律效力。您可以通过中国人保官网（www.epicc.com.cn）对电子保单进行验真。
4. 如需变更保单信息、咨询保险产品相关事宜、理赔信息，请联系中国人保全国统一客户服务及投诉热线：0755-95518。
5. 阅读条款：请确认您已认真阅读《中国人民财产保险股份有限公司意外伤害保险条款》及《中国人民财产保险股份有限公司附加约定或调整责任期间保险条款》，您了解、同意并确认上述条款内容，特别是保险责任及责任免除的部分。作为投保人，您确认对于被保险人具有保险利益。
6. 投保前请确认并了解本产品适用条款，请务必阅读其中的责任免除部分，以下为免责任条款：
  - （一）原因除外  
被保险人因下列原因而导致身故或残疾的，保险人不承担给付保险金责任：

2/8

2.2 你应当通过微保提供或认可的方式使用本服务。你依本协议条款所取得的权利不可转让。

2.3 你不得使用未经微保授权的插件、外挂或第三方工具对本协议项下的服务进行干扰、破坏、修改或施加其他影响。

### 三、【个人信息收集、使用授权与保护】

3.1 为了提升你的投保体验，更好的服务你的投保需求，你同意并授权微保通过以下途径获取、收集你的信息。

（1）收集你留存在保险公司处的信息；

（2）收集你留存在微保合作伙伴、关联公司处的信息；

（3）向行政机关、司法机关查询、留存你的信息；

（4）向合法留存你信息的自然人、法人以及其他组织收集你的信息。

3.2 你同意并授权微保收集你的信息范围包括但不限于：

（1）你在保险公司处留存以及形成的数据和信息；

（2）你在微保关联公司和合作伙伴处中留存以及形成的数据和信息；

（3）你在申请使用微保服务时所提供以及形成的数据和信息；

3/8

（4）与你申请或使用的服务相关的，你留存在其他自然人、法人或其他组织的相关信息。

3.3 你同意并授权微保将你的信息用于如下用途：

（1）为了向你提供微保的产品、服务之用；

（2）帮助我们设计新服务、改善我们现有服务，提升你的保险服务，投保体验之用；

（3）向保险公司提供缔结、履行保险合同及本协议约定服务之用；

（4）向其他相关主体提供以履行本协议约定服务之用；

（5）为了向你提供相关服务，将你的信息分享给联合提供该项服务的关联公司或合作伙伴；

（6）为了更好的为你提供某些产品或服务，可能会向你进行电话呼叫或短信通知。

3.4 为了你的信息安全，微保将严格保护你的信息。

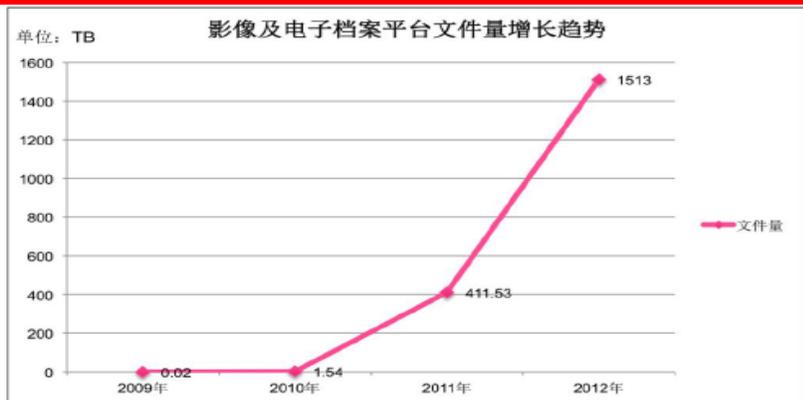
（1）微保将运用各种安全技术和程序建立完善的管理制度来保护你的个人信息，以免遭受未经授权的访问、使用或披露；

（2）微保不会将你的个人信息转移或披露给任何非关联的第三方，除非：

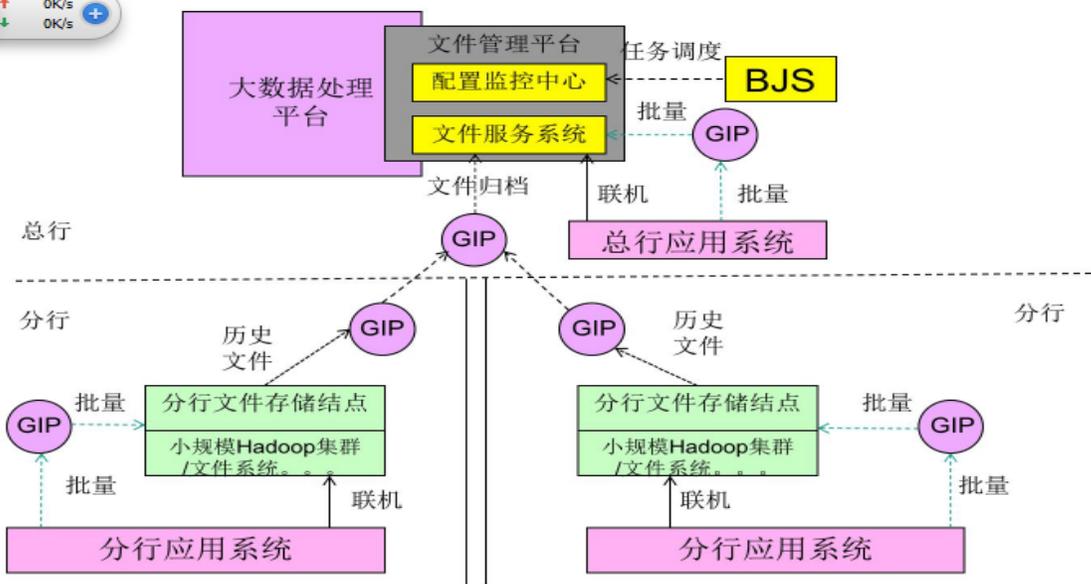
① 相关法律法规或法院、政府机关要求；

② 为完成合并、分立、收购或资产转让而转移。

## (5) 大数据与银行业务优化



32% ↑  
OK/s  
↓  
OK/s



### 客户挑战

- 业务系统中对于历史数据保存和访问的需求越来越多,数据量越来越大,特别是票据影像信息达到了几个PB
- 原有系统不支持PB级历史数据归档存储与数据加工,不支持高并发历史数据的在线查询与批量查询

### 解决方案

- **企业版Hadoop平台**: 高可靠, 高安全, 易管理、易开发
- **主备**大数据业务集群, **100+**大数据节点
- 提供海量数据归档存储和高性能查询能力, 既支持结构化数据的归档查询也支持**非结构化数据**的归档查询

### 客户价值

- 建设容量**6PB**的的票据影像归档系统, 满足全行的票据归档需求
- 在存储空间占用量、历史数据使用便捷性、归档成本之间取得最佳平衡
- 通过历史数据分析为产品设计、更新、监测提供决策支持

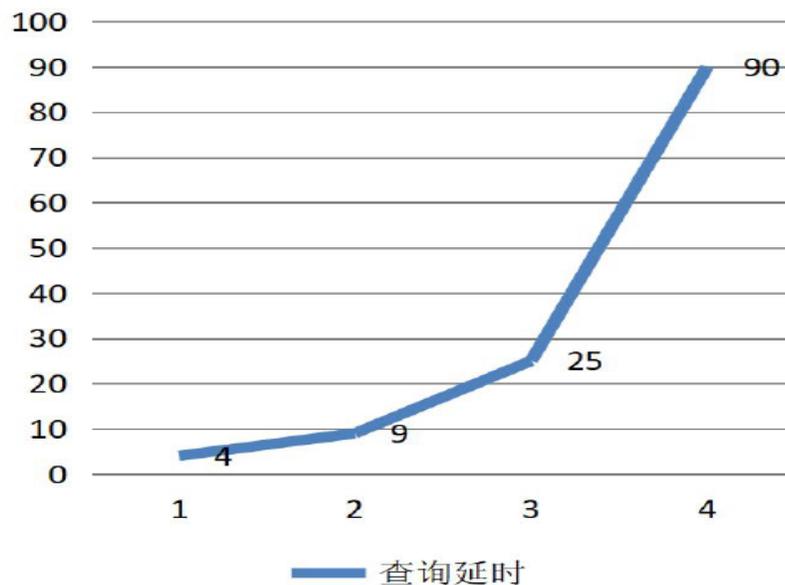
# 民生银行卡部历史工单查询

## 历史数据量（4年）

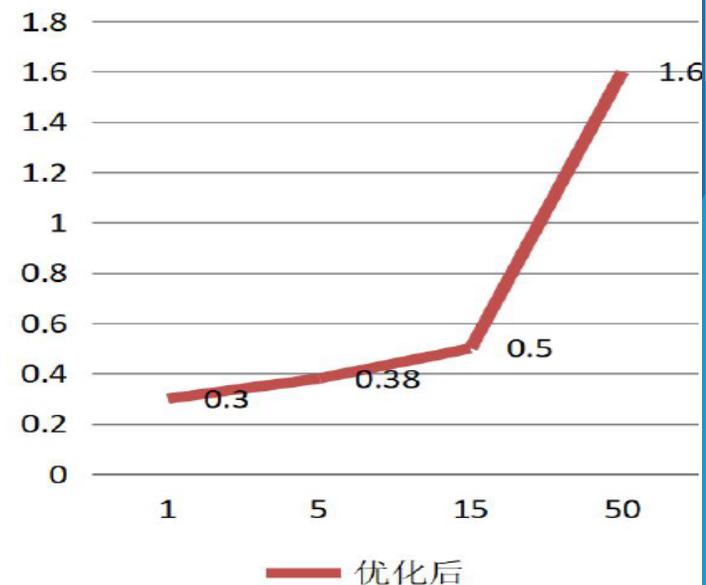
数据表	行数	大小
dds_acct_acct	18124621	2.7G
dds_acct_card	38665299	17.6G
dds_acct_stmt	147632006	69.8G
dds_trans_event	716425258	218.5G
dds_acct_quick_chng	107066634 4	616.5G
tab_info_list	3094	239.4K
合计		925GB

	生产系统(SAS)	TDH
硬件	2x P750小型机 (HA)	8台x86服务器
工单查询延时	最快20分钟	平均4秒
程序	SAS 444行	PL/SQL 108行

### 并发查询延时（秒）



### 并发查询延时（秒）



# 应用7、影视大数据

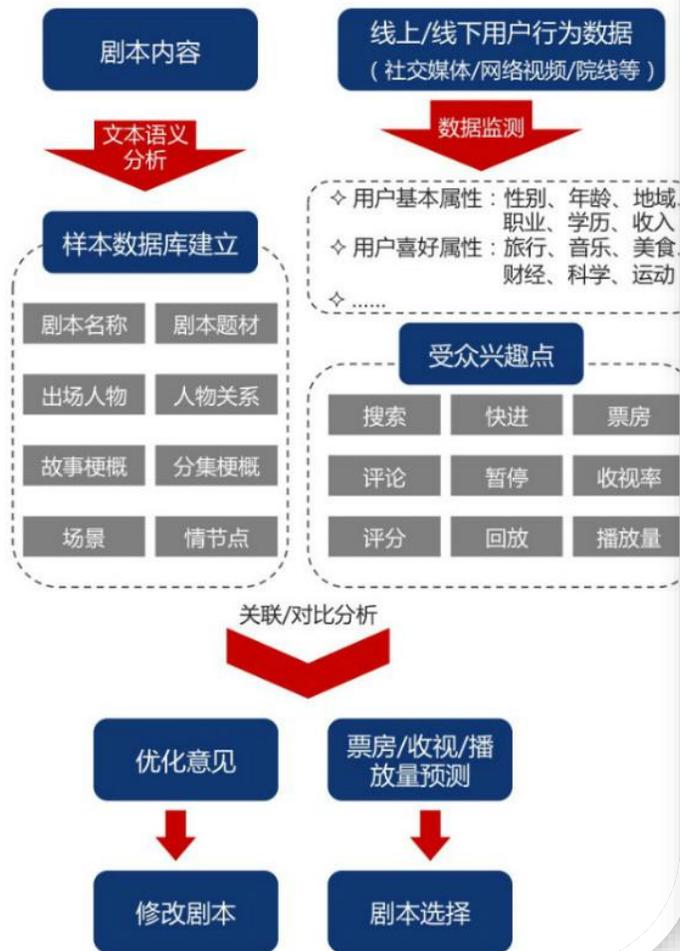


### 2013-2018年中国影视大数据应用服务市场容量



数据来源：数据猿研究院根据公开票房数据、电视剧交易规模数据建立模型估算得出。

## 价值体现1—剧本立项



## 价值体现2—主创选择



## 价值体现3—宣发环节



# 应用8、舆情大数据

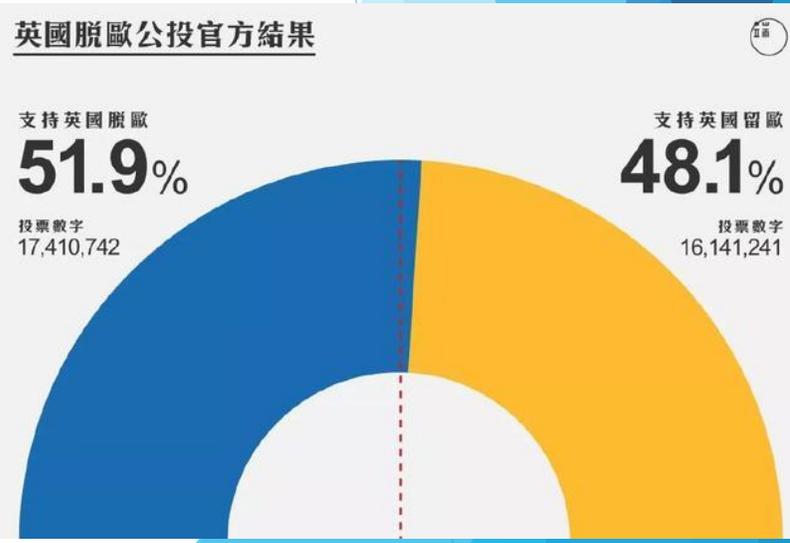
推荐：中青华云大数据平台  
<http://www.1yuqing.com/>

信息技术对政治权力的影响：集中->分散->集中->-----



## Cambridge Analytica (剑桥分析)

“他们针对那些反对英国脱欧的人，给他们一种胜券在握的感觉，就算他们不投票，英国也绝不可能成功脱欧。而对那些支持脱欧的人，它则鼓励每个人走上街头，为自己的选择投票。”



# 应用9、大数据反腐



哈尔滨市人大预算工作委员会原主任朱海的房产多达11套，但是狡猾的他并未把这些房产都写在自己的名下。不过，就在他自以为“安全”之际，哈尔滨市纪委调查人员利用大数据发现，有11套的房的水电气费、物业费都由朱海缴纳！

哈尔滨市公安局交警支队考验处原处长侯宝光任职前后消费水平明显变化：侯任职5年乘航班54次，头等舱占24次，而此前4年乘航班19次，无头等舱；任职后其妻子出国（境）18次，而此前仅出境一次。



# 应用10、工业大数据

## CMS设备状态监测系统

- ▶ 采用有效的检测手段和分析诊断技术，及时、准确的掌握设备运行状态，保证设备的安全、可靠和经济运行。目前采用的状态监测方法有两种：在线设备状态监测和离线设备状态监测。状态检测与故障诊断通俗地讲就是给机器看病。



人：



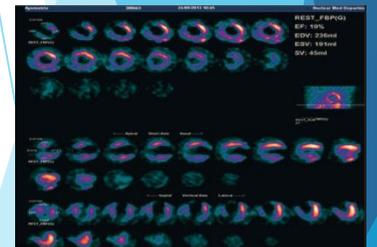
量体温



照X光



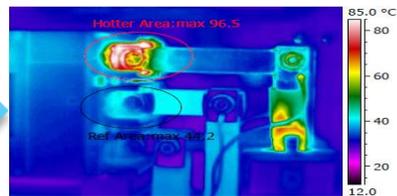
B超



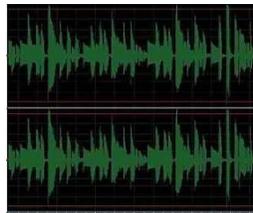
CT



机器：



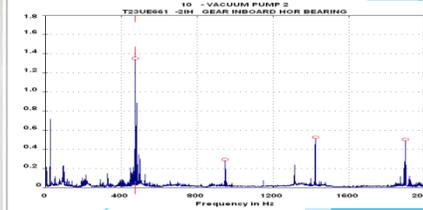
测量温度



测量噪音



测量振动值



油液分析





# 应用11、车辆大数据





无人驾驶清洁车

# 应用12、大数据金融犯罪



# 应用13、与VR、AI技术结合

2019年：腾讯旗下首款AR探索手游《一起来捉妖》



2016年：任天堂游戏《Pokemon Go!》





**RISE Lab**

From **live data** to **real-time decisions**



**AMP Lab**

From **batch data** to **advanced analytics**

技术上：全球顶级实验室Berkeley  
Rise Lab开始转向人工智能算法研究



2017年中国大数据  
人工智能创新创业大赛



2017.04-2017.10

面向全球火热招募中

# 语音识别



## 典型科研院所

斯坦福大学自然语言处理研究小组

卡内基梅隆大学语言技术研究院

约翰霍普金斯大学语言和语音处理研究组

谷歌苏黎世研究院

亚马逊NLP研究组

香港科技大学NLP实验室

中科院计算机所自然语言处理研究组

中科院声学所语言声学与内容理解重点实验室

中科院自动化所模式识别重点室

百度自然语言处理部

中科大人机语音通信研究评测实验室

—36Kr研究院

### 三、上海大数据发展的主要 做法及面临的挑战

## 上海市大数据“五位一体”的顶层规划布局，谋划构建大数据产业发展的良好生态



交易中心



产业基金



产业基地



产业联盟



研究机构

2016年4月1日，上海数据交易中心、上海市大数据产业基地、上海大数据联盟在静安国际大数据论坛上成立。

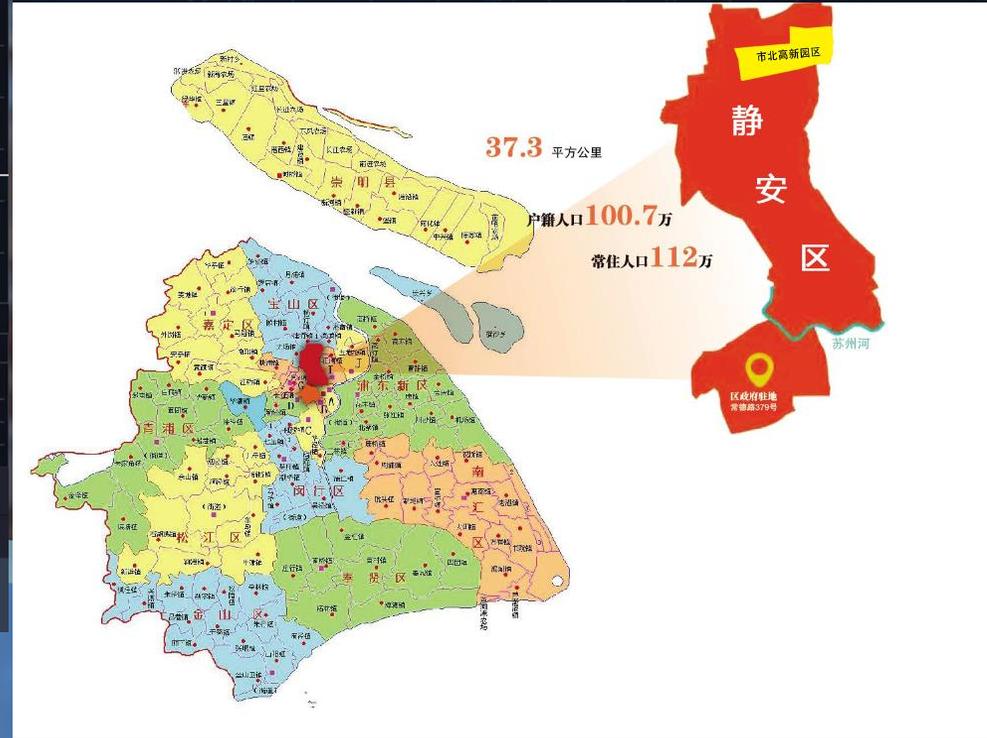
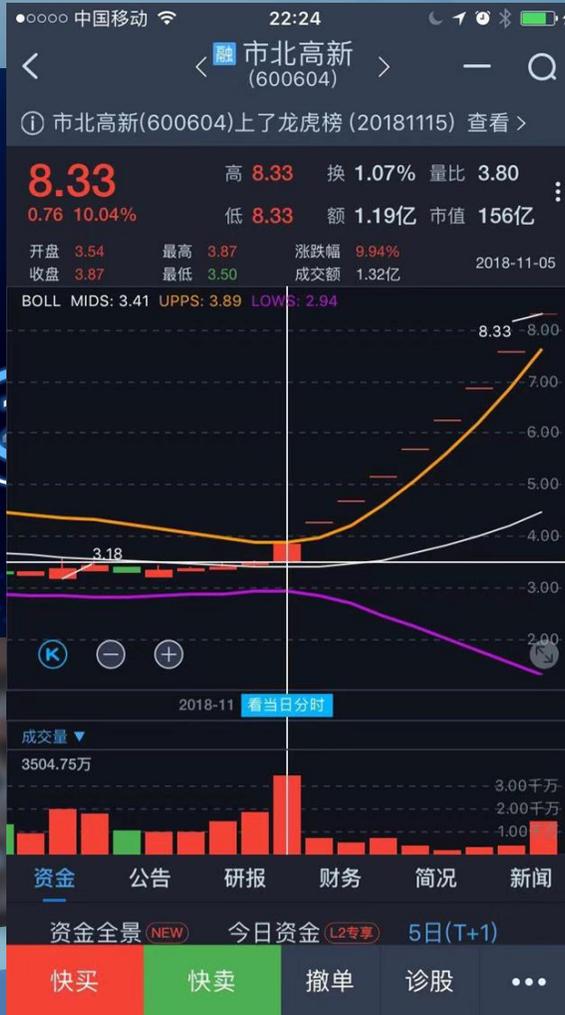
# 上海市大数据产业基地：上海市北高新（集团）有限公司

## “中国大数据产业之都”

市北正式发布《打造中国大数据产业之都三年行动计划》

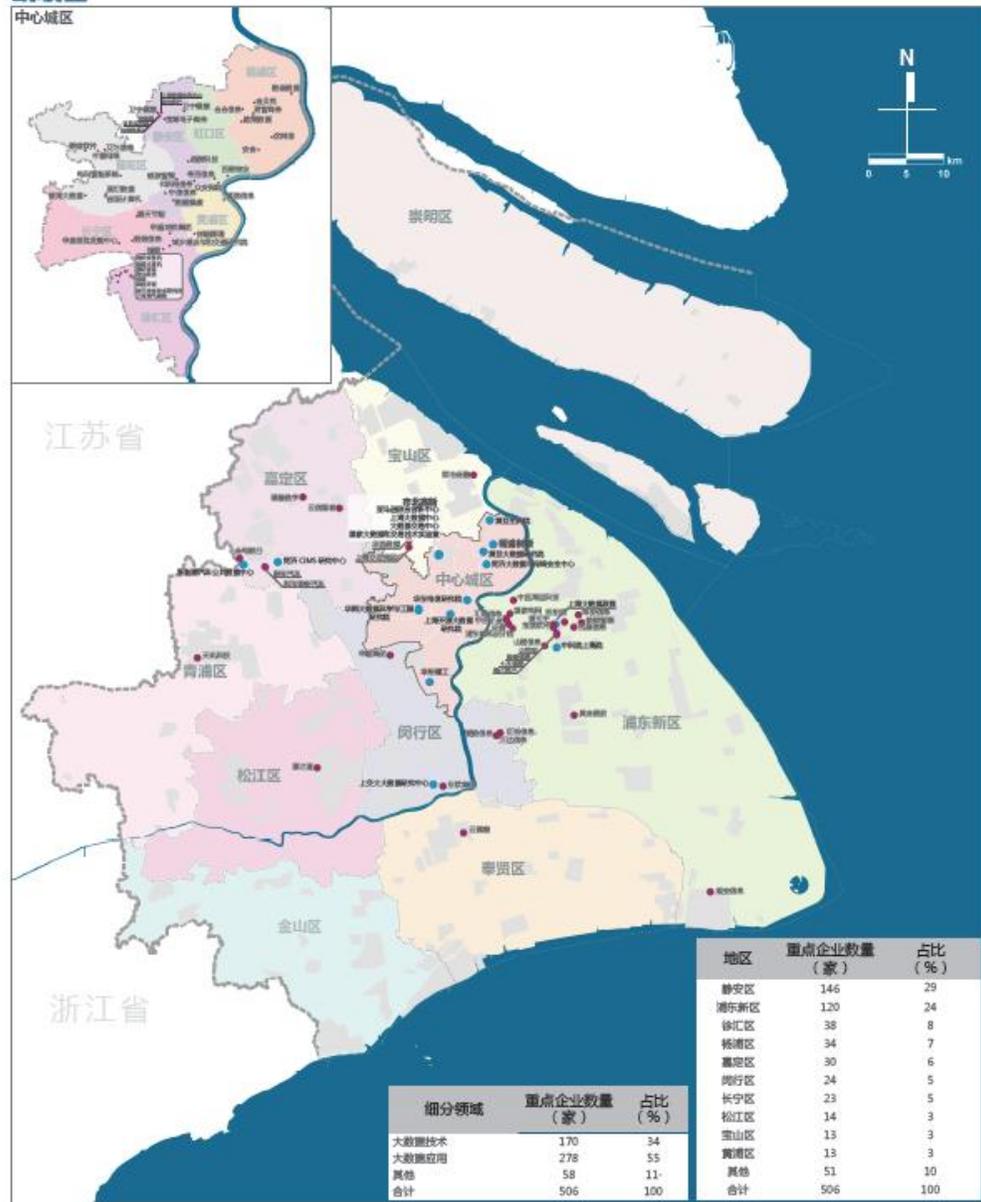
上海全球科创中心的“数据核”

长三角世界级城市群的“数据港”



# 上海市大数据发展地图

## 现状图

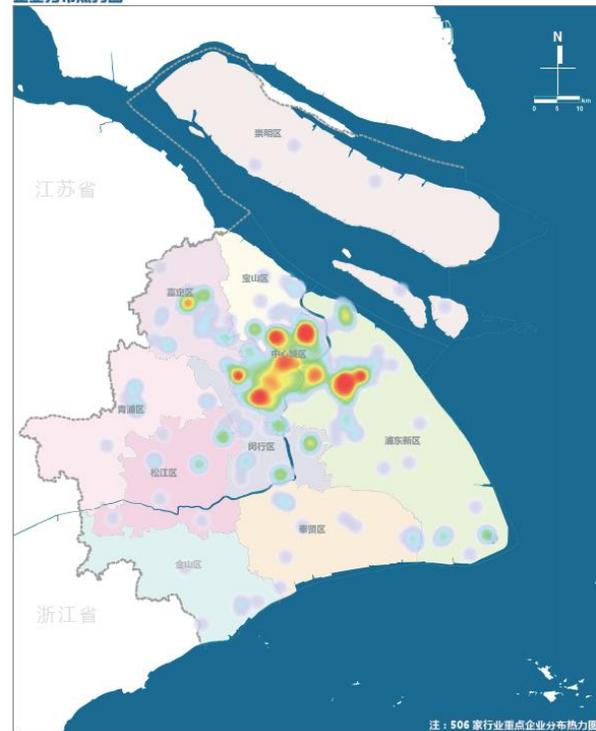


图例  
 ● 创新资源 ● 企业资源  
 - - - 市界 — 区界

上海产业地图集  
 图纸编号 005(1)  
 编制单位 上海市产业发展研究和评估中心

# 上海市大数据发展地图

## 企业分布热力图



图例  
 密集  
 稀疏

上海产业地图集  
 图纸编号 005(1)  
 编制单位 上海市产业发展研究和评估中心

## 规划图: 1+4



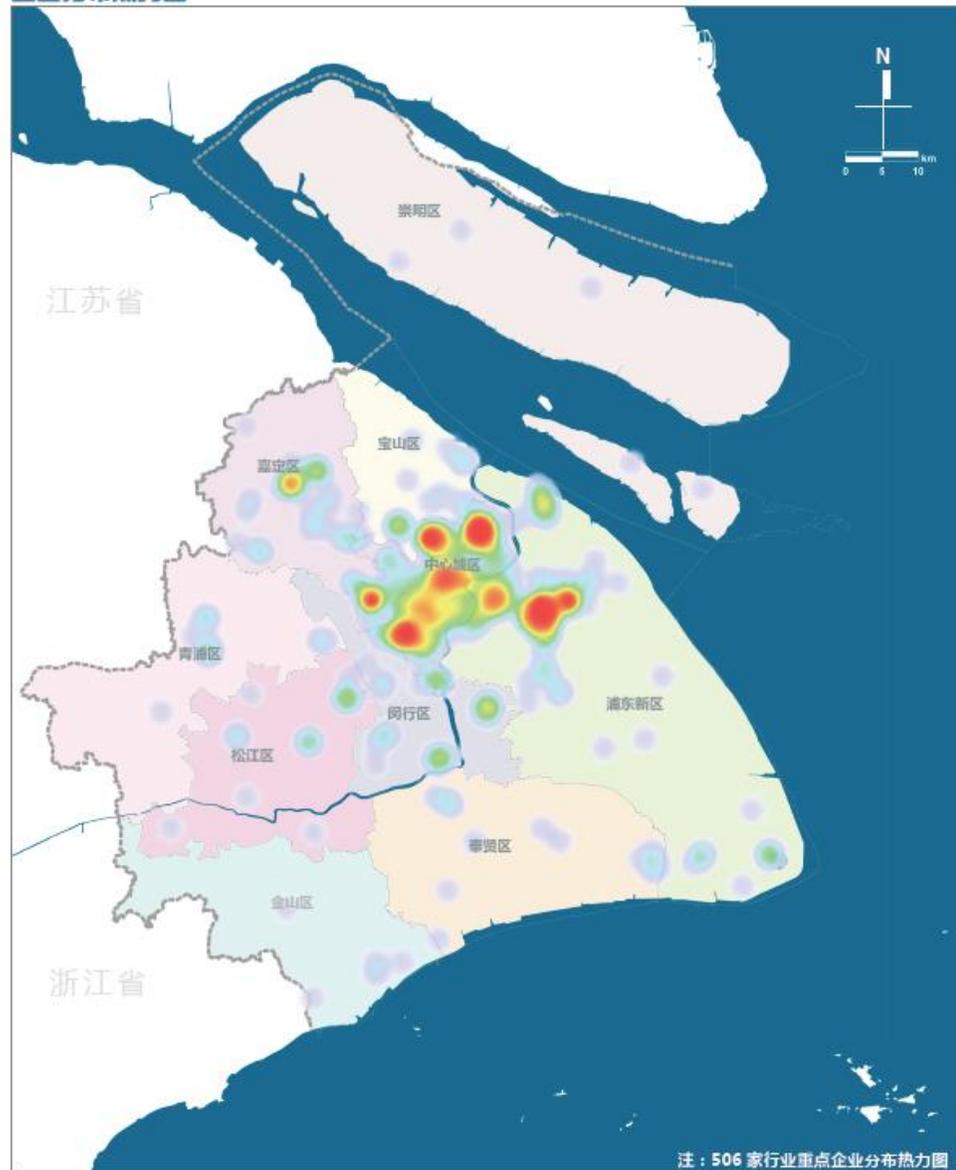
图例  
 ● 核心区域 ● 其他重点区域  
 — 主要河流 - - - 市界 — 区界

大数据“1+4”空间布局  
 1中心 静安区 国家新型工业化大数据示范基地、上海市大数据产业基础、公共数据开放基地  
 4重点 杨浦区 上海市大数据产业基础、公共数据开放基地  
 浦东新区 公共数据开放基地  
 闵行区 公共数据开放基地  
 松江区 长三角大数据产业基地

上海产业地图集  
 图纸编号 005(2)  
 编制单位 上海市产业发展研究和评估中心

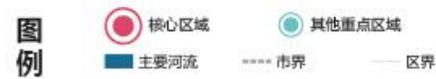
# 上海市大数据发展地图

## 企业分布热力图



上海产业地图集  
图纸编号 005(1)  
编制单位 上海市产业发展研究和评估中心

## 规划图：1+4



上海产业地图集  
图纸编号 005(2)  
编制单位 上海市产业发展研究和评估中心



# 上海大数据联盟

SHANGHAI BIG DATA ALLIANCE



Member logos include:

- 上海华东电信研究院 (East China Institute of Telecommunications)
- 同濟大學 (TONGJI UNIVERSITY)
- 辰智咨询 (CHENZHI CONSULTING)
- 梅斯医药 (MedSci China)
- 航天机电
- 云健康 (CLOUD HEALTH)
- 中科曙光 (Sugon)
- 信数金服 (不止于数据)
- 上海嘉值 (VAS Technology)
- 华院数据
- 晶赞科技 (ZAMPLUS TECHNOLOGY)
- MINIDATA (麦达数字)
- GLORYSCIENCE (上海国南信息技术有限公司)
- 华大医学 (BGI·DX)
- 东方申信 (Oriental Info.Tech.)
- 华东理工大学 (EAST CHINA UNIVERSITY OF SCIENCE AND TECHNOLOGY)
- 上海生物芯片 (inspur 浪潮)
- 通付盾 (PayEgis)
- seassoon
- 市北高新 (SHIBEI HI-TECH)
- 上海超级计算中心 (Shanghai Supercomputer Center)
- CETC
- 上海交通大学 (SHANGHAI JIAO TONG UNIVERSITY)
- 上海亚太
- 经尔纬数据技术 (Matrix United Technologies)
- BOCOM (博康智能)
- 万达信息
- 华谊集团 (HUAYI)
- 东方财富
- 海譽知科技 (Hiknowledge Technology)
- 天玑科技 (Pioneer in IT Service)
- E·WE 亿微 (上海邮电设计咨询研究院)
- 鹏博士集团 (DR.PENG GROUP)
- 申能集团
- 上海产业技术研究院 (Shanghai Industrial Technology Institute)
- 复旦大学 (FUDAN UNIVERSITY)
- 中国电信 (CHINA TELECOM)
- 世界触手可及
- Double Click (world trade)
- China unicom (中国联通)
- Yun Goal
- 畅享网 (服务于中国管理人)

上海大数据联盟会员已超过600家



1.举办了上海地区唯一的全行业大数据产业峰会。该峰会于2016年12月1日-2日、2017年11月30日-12月1日举办。参会人数超过2000人。



2.创立了国内首个大数据行业应用系列论坛——魔方大数据系列论坛。东方卫视、第一财经等多家电视台多次报道。已举办26期。



3、合作举办了国内首个大数据领域的人工智能大赛——BOT大赛。2016年7月启动，11月决赛；2017年3月启动了上海大数据人工智能大赛。



4、上海大数据联盟静安大数据创新加速器，2018年4月3日正式挂牌。11家入驻企业，涵盖医疗、精准营销、能源、地理、金融等各方面大数据的应用。



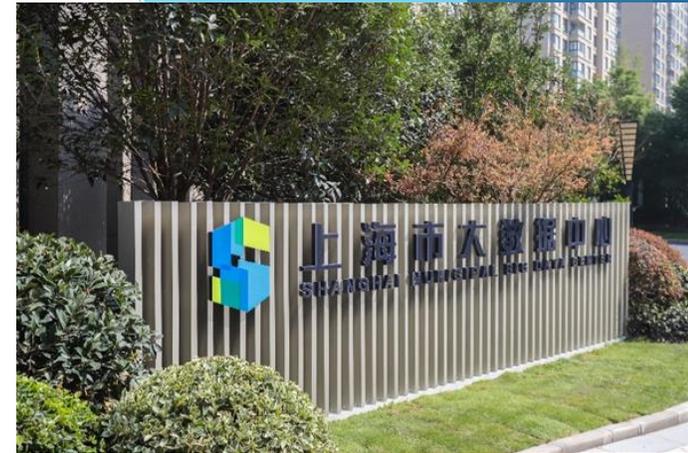
2018年7月16日下午，长三角大数据产业合作联席会议第一次工作会议在沪顺利召开。



附：长三角大数据产业合作联席会议成员单位  
(排名不分先后)

- 1、上海大数据联盟
- 2、复旦大学智慧城市研究中心
- 3、上海浦东数据工程中心
- 4、上海市大数据产业基地（市北高新）
- 5、上海大数据应用创新中心
- 6、上海云部落TMT产业园
- 7、紫竹园区、中国(上海)网络视听产业基地
- 8、闵行区奇士集群产业园（大数据基地）
- 9、零号湾全球创新创业集聚区
- 10、江苏大数据联盟
- 11、盐城大数据产业园
- 12、溧阳市人民政府
- 13、泰州数据园管委会
- 14、上海市北高新集团（南通）有限公司
- 15、苏州市人工智能和大数据产业联盟
- 16、苏州市相城区人工智能和大数据产业联盟
- 17、昆山市大数据产业联盟
- 18、浙江省大数据应用产业技术联盟
- 19、杭州国信大数据应用研究院
- 20、湖州市大数据管理中心
- 21、浙江省互联网协会
- 22、安徽省大数据产业联盟
- 23、合肥上海产业园
- 24、合肥蜀山经济开发区管理委员会
- 25、淮南高新区管委会

2018年4月12日，上海市大数据中心正式成立。





上海开放数据创新应用大赛



2017年8月，市民云APP入驻Apple Store，正式对外提供to C数据服务。



自2015年开始，举办SODA上海开放数据创新应用大赛。这是国内第一个大规模的政府数据开放创新大赛。

《上海市公共数据开放管理办法》将于2019年发布。

# 2017年度上海大数据发展十大事件

1. 大数据流通与交易技术国家工程实验室在沪成立
2. 上海大数据社会治理创新试点工程项目启动实施
3. 上海大数据应用创新中心获批建设
4. 上海市政务数据资源开放总量突破**1500**项
5. 大数据国际合作交流系列活动成功举办
6. **2017**上海静安国际大数据论坛顺利举行
7. 第三届上海开放数据创新应用大赛成功举办
8. 长三角江浙沪地区大数据产业联盟达成战略合作
9. 上海数据交易中心商业数据交易总量突破**10**亿次
10. 上海大数据应用展示中心累计接待参访人员达**1**万人次

# 2018年度上海大数据发展十件大事

- 1、上海市大数据中心揭牌成立
- 2、《上海市公共数据和一网通办管理办法》正式实施
- 3、上海市获批成为首批国家公共信息资源开放试点城市
- 4、上海静安成功获批首批国家新型工业化（大数据）示范基地
- 5、上海静安“151工程”获批成为国家发改委数字经济试点重大工程
- 6、上海市首次授牌启动七家大数据联合创新实验室试点建设
- 7、十个项目入选国家工信部2018年度大数据产业发展试点示范项目
- 8、首届长三角数据智能合作（上海）峰会隆重举行
- 9、“数联长三角、众创新生活”开放数据创新应用大赛成功举办
- 10、“智能城市，数聚有为”上海静安国际大数据论坛顺利举行

# 大数据产业发展中的主要问题

## 数据安全：

- 1、“斯诺登事件”之后，各国注重数据主权，数据跨境流动受到限制；
- 2、基础设施安全问题。作为大数据汇集的主要载体和基础设施，云计算为大数据提供了存储场所、访问通道、虚拟化的数据处理空间。因此，云平台中存储数据的安全问题也成为阻碍大数据发展的主要因素。
- 3、病毒频发。2017年5月，勒索病毒爆发。被感染的用户必须在7天内交纳300美元等价的比特币作为赎金，否则电脑数据将被全部删除且无法修复。



返回 分布式查询-

旅馆住宿	湖南	广州七	有限公	2016-	413	2706	201	230
旅馆住宿	广西	有限公	司	2016-08	1-52-15	8643	2016-0	32-04
旅馆住宿	广西	有限公	司	2016-09	1-20-51	8630	2016-0	28-41
旅馆住宿	北京	北	京	2015	119	220	201	200
旅馆住宿	北京	北	京	2013	014	3021	201	900
旅馆住宿	北京	北	京	2015	119	220	201	943
旅馆住宿	山西	阳城	实业	2011	623	3012	201	328
旅馆住宿	山西	太原	务有限	2012	023	8305	201	324
旅馆住宿	山西	店		2012	714	8212	201	107
旅馆住宿	山西	店		2013	305	223	201	836
旅馆住宿	山西	店		2015	933	8409	201	332
旅馆住宿	山西	店		2016	156	308	201	633
旅馆住宿	安徽	黄山	旅(黄山	2014	026	8266	201	249
常住人口	吉林							
常住人口	公安部							
暂住人口	山西	P140	9120911					
暂住人口	山西	1401	911251	20	2	20909914		
暂住人口	山西	D140	9127284	20	8			
网吧上网	北京	北京	服务有	20160	4259	19000101000000	海淀区中	西侧社
网吧上网	北京	北京	服务有	20160	4259	20160301171550	海淀区中	西侧社
网吧上网	北京	北京	服务有	20150	1507	20150905182138	海淀区中	西侧社
网吧上网	北京	北京	服务有	20150	1507	19000101000000	海淀区中	西侧社
网吧上网	北京	北京	服务有	20150	1733	19000101000000	海淀区中	西侧社
网吧上网	北京	北京	服务有	20150	1733	20150905160116	海淀区中	西侧社
网吧上网	北京	北京	服务有	20150	4644	19000101000000	海淀区中	西侧社
网吧上网	北京	北京	服务有	20150	4644	20150905131733	海淀区中	西侧社
网吧上网	北京	北京	服务有	20091	4		海淀区信	28号C5
网吧上网	北京	北京	服务有	20091	05		西城区裕	14号楼
网吧上网	北京	北京	服务有	20091	1255	20091113093505	西城区裕	14号楼
网吧上网	北京	北京	服务有	20091	5335	20091113080352	西城区裕	14号楼

【旅馆住宿】	【广东】	厂家酒店公寓	广州市越秀区...	2016		0057	20161031	详细 同住宿
【旅馆住宿】	【重庆】	如家连锁酒店	沙坪坝区沙杨...	20160		428	201608081	详细 同住宿
【旅馆住宿】	【重庆】	如家连锁酒店	沙坪坝区沙杨...	20150		522	201503021	详细 同住宿
【旅馆住宿】	【重庆】	诗苑大酒店	奉节县夔洲路...	20150		706	201501221	详细 同住宿
【旅馆住宿】	【重庆】	戴斯大酒店	南川区南城街...	20150		1202	201501201	详细 同住宿
【旅馆住宿】	【陕西】	安馨宾馆	西安是新城...	2011		310		详细 同住宿

GO 共1页

南方都市报

# 隐私数据买卖泛滥成灾!

旅客编号: 4401040762201610300048 姓: [REDACTED]

民族: 汉族 性别: 男性

出生日期: [REDACTED] (24岁) 证件类型: 11

证件号: [REDACTED] 住址省市区: [REDACTED]

住址址: [REDACTED] 入住时间: 2016 [REDACTED]

入住房号: [REDACTED] 退房时间: 2016 [REDACTED]

旅馆编码: 4401040762 旅馆名称: 厂家酒店公寓

旅馆地址: 广州市越秀区东风东路753号天誉大厦东塔二楼 旅馆地址区划: 广东省广州市

南方都市报

微信(2)

开着机是吧

应该是, 刚刚还发过微信

嗯

1320... 75。 23.125032,113.309044

南方都市报 奥一网 oeeee.com

# 2016年11月7日通过， 自2017年6月1日起正式实施。

## 法律诉讼



### 中华人民共和国 网络安全法

含草案说明

中国法制出版社

第四十条 【网络运营者应当对其收集的用户信息严格保密，并建立健全用户信息保护制度。】

第四十一条 【网络运营者收集、使用个人信息，应当遵循合法、正当、必要的原则，公开收集、使用规则，明示收集、使用信息的目的、方式和范围，并经被收集者同意。网络运营者不得收集与其提供的服务无关的个人信息，不得违反法律、行政法规的规定和双方的约定收集、使用个人信息，并应当依照法律、行政法规的规定和与用户的约定，处理其保存的个人信息。】

第四十二条 【网络运营者不得泄露、篡改、毁损其收集的个人信息；未经被收集者同意，不得向他人提供个人信息。但是，经过处理无法识别特定个人且不能复原的除外。

网络运营者应当采取技术措施和其他必要措施，确保其收集的个人信息安全，防止信息泄露、毁损、丢失。在发生或者可能发生个人信息泄露、毁损、丢失的情况时，应当立即采取补救措施，按照规定及时告知用户并向有关主管部门报告。】

2018年5月25日，GDPR（General Data Protection Regulation一般数据保护条例）正式在欧盟成员国内执行



**数据收集者的操作规范。**针对数据收集者，不能用隐藏默认的方式获取用户许可，必须提前进行明确的提示与询问，获得允许后才可以获取使用用户数据；收集之后还需要为用户提供查看收集数据概览及用途，还必须为用户提供删除功能。

**用户对自己数据完全自主的权利。**在用户权利方面，用户对自己的数据拥有完全的所有权，即便同意收集方收集，也可以随时查看撤回删除相关协议，在用户撤回删除相关授权后，数据收集者必须立即将相关数据进行匿名化处理。

# 本人发表的各类大数据相关论文和内参

- (1) 《互联网治理中的三重性分析》，发表在《中国浦东干部学院学报》，2014年7月，（人大复印报刊资料全文转载，2014年第11期D01《公共行政》）；
- (2) 《大数据拓展国家治理空间》，发表在《社会观察》，2014年8月；
- (3) 《大数据背景下我国金融反恐工作的几点思考》，发表在中国浦东干部学院《领导参考》，2014年8月；
- (4) 《实现政府大数据价值的几点思考》，发表在中国浦东干部学院《研究咨询》，2014年11月；
- (5) 《大数据推动我国法治建设进程》，发表在《社会观察》，2014年11月，被“求是理论网”全文转载；
- (6) 《上海自贸区建立离岸互联网数据中心的思考》，发表在中国浦东干部学院《领导参考》，2015年9月；
- (7) 《Collaborating CPU and GPU for the electromagnetic simulations with the FDTD algorithm》，发表在《Concurrency & Computation Practice & Experience》，2016年4月；
- (8) 《基因大数据产业离腾飞还有多远？》，发表在《浙商管理评论》，2016年5月；
- (9) 《大数据产业发展的国际经验》，发表在《张江科技评论》，2017年6月；

# 上海大数据联盟



欢迎加入我们  
微信公众号扫一扫

