



# 大数据推动政府治理变革

何宝宏 中国信息通信研究院

2017年6月8日

- 一、大数据概论与政策解读**
- 二、大数据产业生态与应用实践案例**
- 三、大数据推动政府治理变革**
- 四、大数据合规性与《网络安全法》**

---

# 一、大数据概论与政策解读

- 大数据的概念
- 全球大数据发展形势
- 我国大数据发展政策
- 地方大数据发展情况

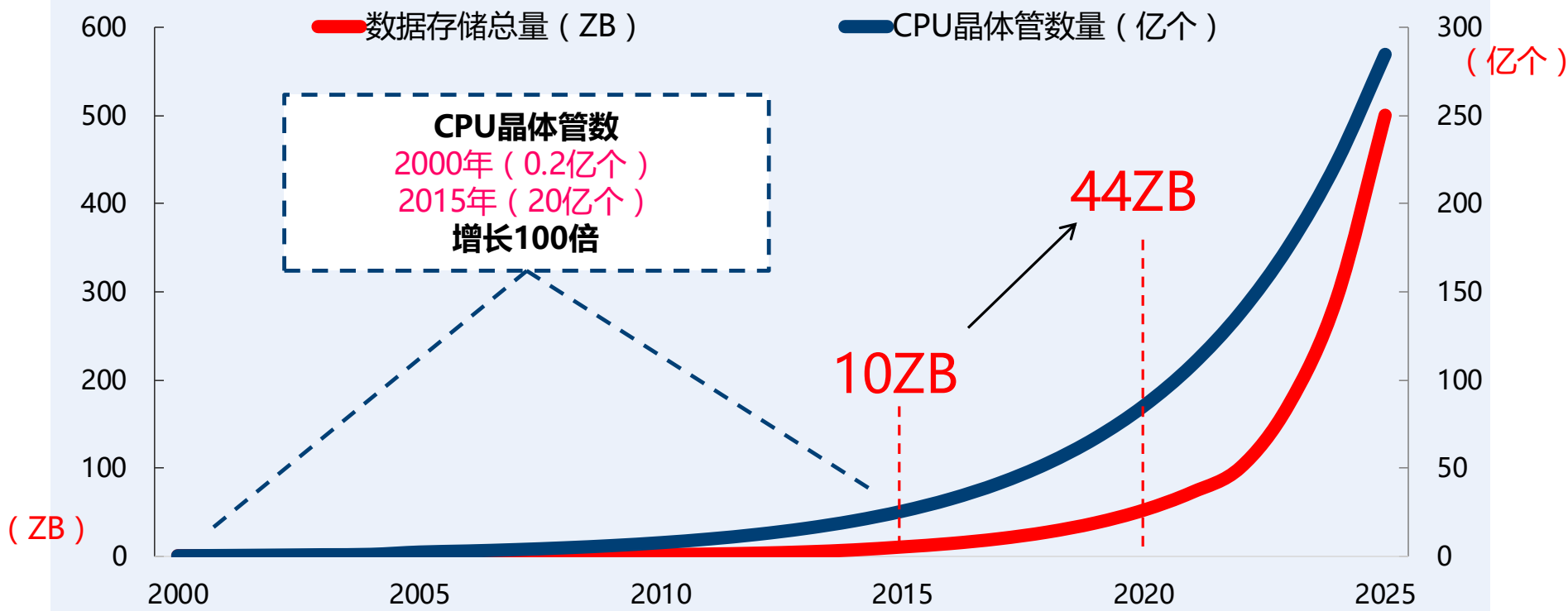
---

- **大数据的概念**

# 万物互联时代 数据爆炸式增长

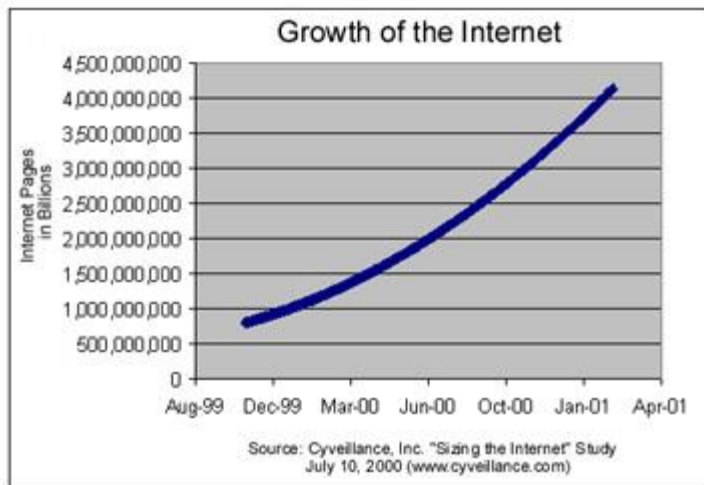
全球互联网步入泛在普及、深度融合、变革创新、引领转型的新阶段

- 数据存储量指数级增长。2015年，10ZB，2020年，达44ZB，是15年的4倍多
- 计算机处理能力按摩尔定律增长。2000-2015年CPU晶体管数量增加100倍
- 存储设备价格大幅下降。2000-2015年硬盘价格降低为1/200



数据来源：国际数据公司，中国信息通信研究院

# 大数据时代的起源



1999-2000网页数量从5亿增长到40亿，每天新增700万！PB级非结构化数据

互联网内容的暴增是第一推动力量

**大数据起源**

2000年前后，基于海量数据分析的搜索引擎逐步发展

**大数据时代**

2006后，电商、广告、SNS应用，2011大数据提出

**突变**

分布式文件系统  
分布式并行计算  
分布式数据库

**数据耦合**

1950，计算机诞生，**数据与应用捆绑**，都存储在文件中

**数据库**

1960，**数据与应用分离，事务处理**  
数据库技术发展

**数据分析**

1990，提出  
**数据挖掘和商业智能**

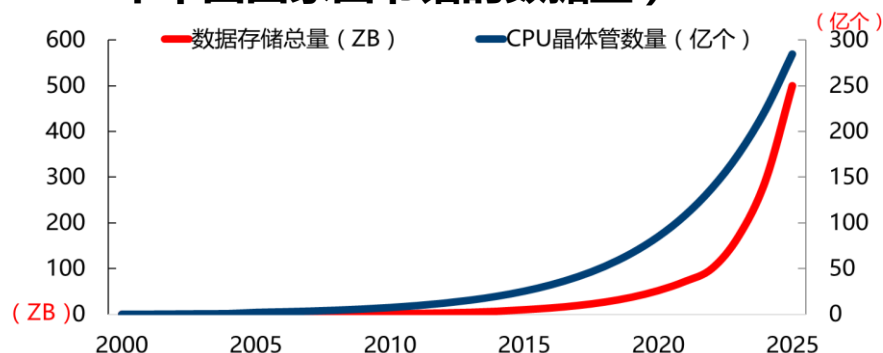
今天：随着应用领域扩展，技术也在不断演进，为更广泛应用提供条件

# 大数据的特征

- 人类社会已经进入大数据时代，全球数据量每18个月翻一番
- “大数据”定义目前尚无统一定论，业界多数认为大数据是**4V特征的数据集合**

## Volume数据量大

2020年全球数据总量将达44ZB  
(1ZB=1万G比特字节，相当于776万个中国国家图书馆的数据量)



## Variety类型复杂

2020年我国非结构化数据占比达到90%



## Velocity速度快

实时快速数据处理、实时结果导向



## Value价值

运用非常规工具完成数据处理，产生巨大价值



---

- **全球大数据发展形势**



# 各主要国家积极布局大数据

——发达国家均高度重视大数据在促进经济社会创新变革、提升国家竞争力中的重要作用，出台国家战略，落实配套措施，系统推动大数据发展



美国认为大数据事关国家竞争力，多措并举推动发展

开放数据、资助研发、带头应用、隐私保护



欧盟：2014年  
出台数据驱动经济战略



澳大利亚：2013年  
发布了公共服务大数据战略



新加坡：建设数据地平线平台（RAHS），数据治国

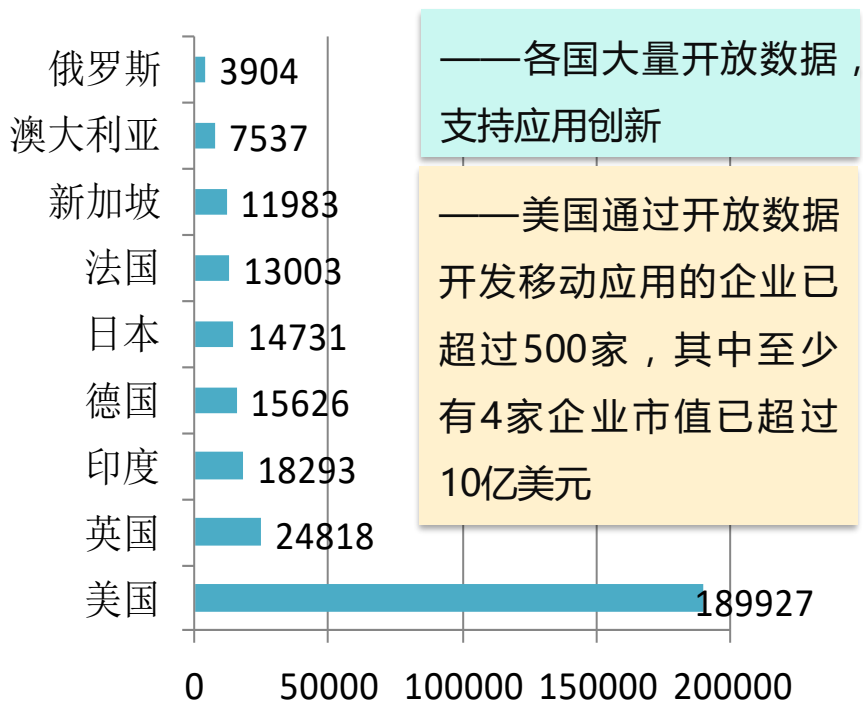


日本发布  
《创建最尖端IT国家宣言》

# 发达国家推进大数据发展的主要方向（1）

——发达国家从推进数据开放、加强技术研发、构建产业生态和完善法律法规四方面促进大数据发展

**1、数据开放：**实施数据开放的国家达97个，支持大批创新应用和创业



各国数据开放网站开放数据集量（个）（截至2015.10.29）

**2、技术研发：**加大投入支持基础、核心、前沿、公共应用研发



# 发达国家推进大数据发展的主要方向（2）

## 3、产业生态：强化公私合作和跨部门协作，促进形成大数据生态体系

美



研发支持中注重多部门协同，强调与大学、研究机构和企业协同

## 4、法律法规：立法探索，以应对个人隐私保护、知识产权保护等新问题

——发布白皮书，对保护个人隐私和公平权利提出建议

欧



与行业协会签署PPP协议，制定政策、运作项目，引导产业投资、研发创新

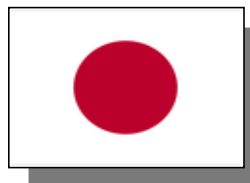
——探索解决个人数据保护、文本挖掘著作权保护、数据所有权建立等问题

澳



通过公私合作促进公共服务大数据发展，扶持商业创新

日



修订《个人信息保护法》，允许企业向第三方出售充分匿名化数据

---

- **我国大数据发展政策**

# 习近平：建设全国一体化的国家大数据中心



以推行电子政务、建设新型智慧城市等为抓手，以数据集中和共享为途径，建设全国一体化的国家大数据中心，推进技术融合、业务融合、数据融合，实现跨层级、跨地域、跨系统、跨部门、跨业务的协同管理和服务。

——2016年10月9日，习近平主持第三十六次中共中央政治局集体学习时发表讲话

要着力推动互联网与实体经济深度融合发展，以信息流带动技术流、资金流、人才流、物资流，促进资源配置优化，促进全要素生产率提升，为推动创新发展、转变经济发展方式、调整经济结构发挥积极作用。

——2016年4月19日，习近平主持召开网络安全和信息化工作座谈会并发表讲话

# 李克强：大数据是“钻石矿”



要进一步推进“互联网+”行动，广泛运用物联网、大数据、云计算等新一代信息技术，促进不同领域融合发展，催生更多的新产业、新业态、新模式，推出更加符合市场需要的新产品和服务。

——2016年6月27日，李克强总理在第十届夏季达沃斯论坛上的演讲

大数据等新一代互联网技术深刻改变了世界，也让各国站在科技革命的同一起跑线上。中国曾屡次与世界科技革命失之交臂，今天要把握这一历史机遇，抢占先机，赢得未来。

——2016年5月25日，李克强总理在贵阳出席中国大数据产业峰会上的演讲

深入实施《中国制造2025》，加快大数据、云计算、物联网应用，以新技术新业态新模式，推动传统产业生产、管理和营销模式变革。”。

——2017年3月5日，李克强总理在十二届全国人大五次会议上作政府工作报告，这是“大数据”第四次写入政府工作报告

# 我国实施国家大数据战略

## 《中华人民共和国国民经济和社会发展第十三个五年规划纲要》第二十七章：实施国家大数据战略

### 第一节 加快政府数据开放共享

- 全面推进重点领域大数据高效采集、有效整合。
- 加快建设国家政府数据统一开放平台。
- 制定政府数据共享开放目录，依法推进数据资源向社会开放。

### 第二节 促进大数据产业健康发展

- 深化大数据在各行业的创新应用，探索与传统产业协同发展新业态新模式，加快完善大数据产业链。
- 加快海量数据采集、存储、清洗、分析发掘、可视化、安全与隐私保护等领域关键技术攻关。
- 促进大数据软硬件产品发展。
- 完善大数据产业公共服务支撑体系和生态体系，加强标准体系和质量技术基础建设。

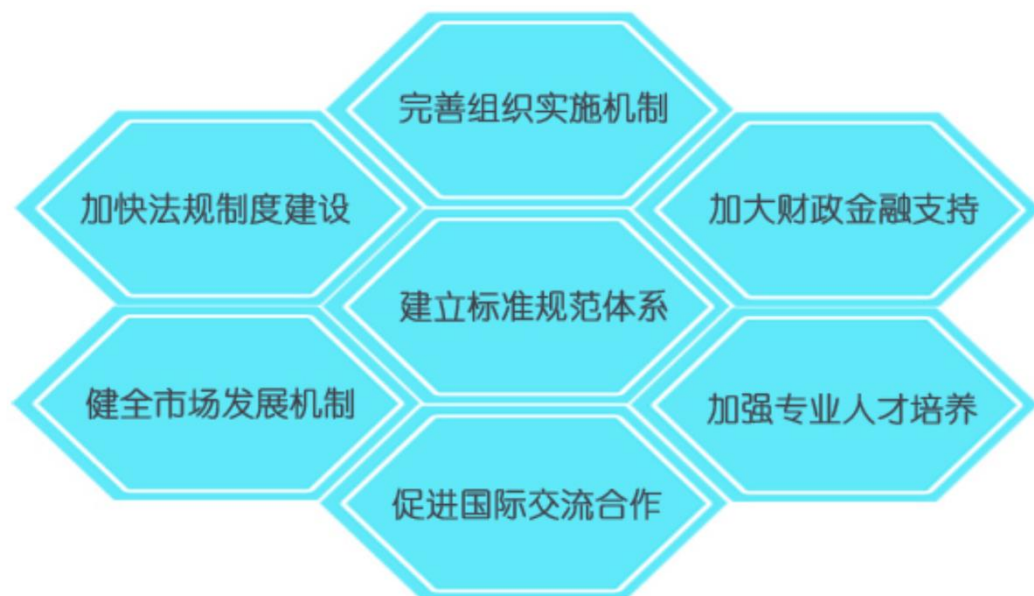




# 国家顶层设计：《促进大数据发展行动纲要》

2015年8月31日，国务院正式印发《促进大数据发展行动纲要》（国发〔2015〕50号），从总体目标、主要任务、政策机制等方面全面部署了我国大数据的发展。为国务院各部委、地方政府制订大数据行业发展规划提供了指导思想和重要依据。

——中国信息通信研究院是《行动纲要》出台的核心支撑单位





# 《促进大数据发展行动纲要》 内容介绍

## 《促进大数据发展行动纲要》

一个目标

全面推动我国大数据发展和应用，加快建设数据强国

三大内容

加快政府数据开放  
推动资源整合  
提升治理能力

推动产业创新发展  
培育新业态  
助力经济转型

健全安全保障体系  
强化安全支撑  
提高管理水平

十项工程

政府  
数据  
资源  
共享  
开放  
工程

国家  
数据  
资源  
统筹  
发展  
工程

政府  
治理  
数据  
工程

公共  
服务  
数据  
工程

工业  
和新  
兴产  
业大  
数据  
工程

现代  
农业  
数据  
工程

万众  
创新  
数据  
工程

关键  
技术  
及产  
品研  
发与  
产业  
化工  
程

大数  
据产  
业支  
撑能  
力提  
升工  
程

网络  
和大  
数据  
安全  
保障  
工程

七条措施

完善组织实施机制

加快法规制度建设

健全市场发展机制

建立标准规范体系

加大财政金融支持

加强专业人才培养

促进国际交流合作

# 工信部：大数据产业发展规划（2016-2020）

为全面部署“十三五”时期大数据产业发展工作，加快建设数据强国，为实现制造强国和网络强国提供强大的产业支撑，工业和信息化部于2016年12月公布了《大数据产业发展规划（2016-2020年）》。规划提出了我国大数据产业发展总体目标及五个分目标为：

**到2020年，技术先进、应用繁荣、保障有力的大数据产业体系基本形成。大数据相关产品和服务业务收入突破1万亿元，年均复合增长率保持30%左右，加快建设数据强国，为实现制造强国和网络强国提供强大的产业支撑。**

**——中国信息通信研究院是《规划》编制的牵头支撑单位**

**技术产品先进可控**

**应用能力显著增强**

**生态体系繁荣发展**

**支撑能力不断增强**

**数据安全保障有力**

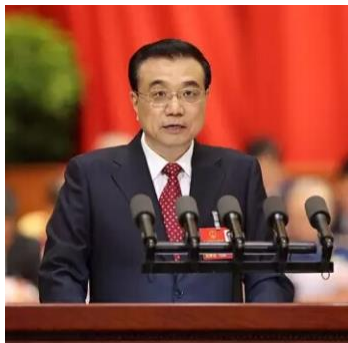
# 大数据产业发展规划：主要内容



---

- **地方大数据发展情况**

# 国家大数据综合试验区：大数据发展需要示范引领



“坚持创新驱动发展，加快大数据部署，深化大数据应用，已成为稳增长、促改革、调结构、惠民生和推动政府治理能力现代化的**内在需要和必然选择**。”

“开展区域试点，推进**贵州等大数据综合试验区**建设，促进区域性大数据基础设施的整合和数据资源的汇聚应用。”

——国务院关于印发促进大数据发展行动纲要的通知  
(国发〔2015〕50号)



中华人民共和国国家发展和改革委员会  
National Development and Reform Commission



中华人民共和国工业和信息化部  
Ministry of Industry and Information Technology of the People's Republic of China



中共中央网络安全和信息化领导小组办公室  
Office of the Central Leading Group for Cyberspace Affairs

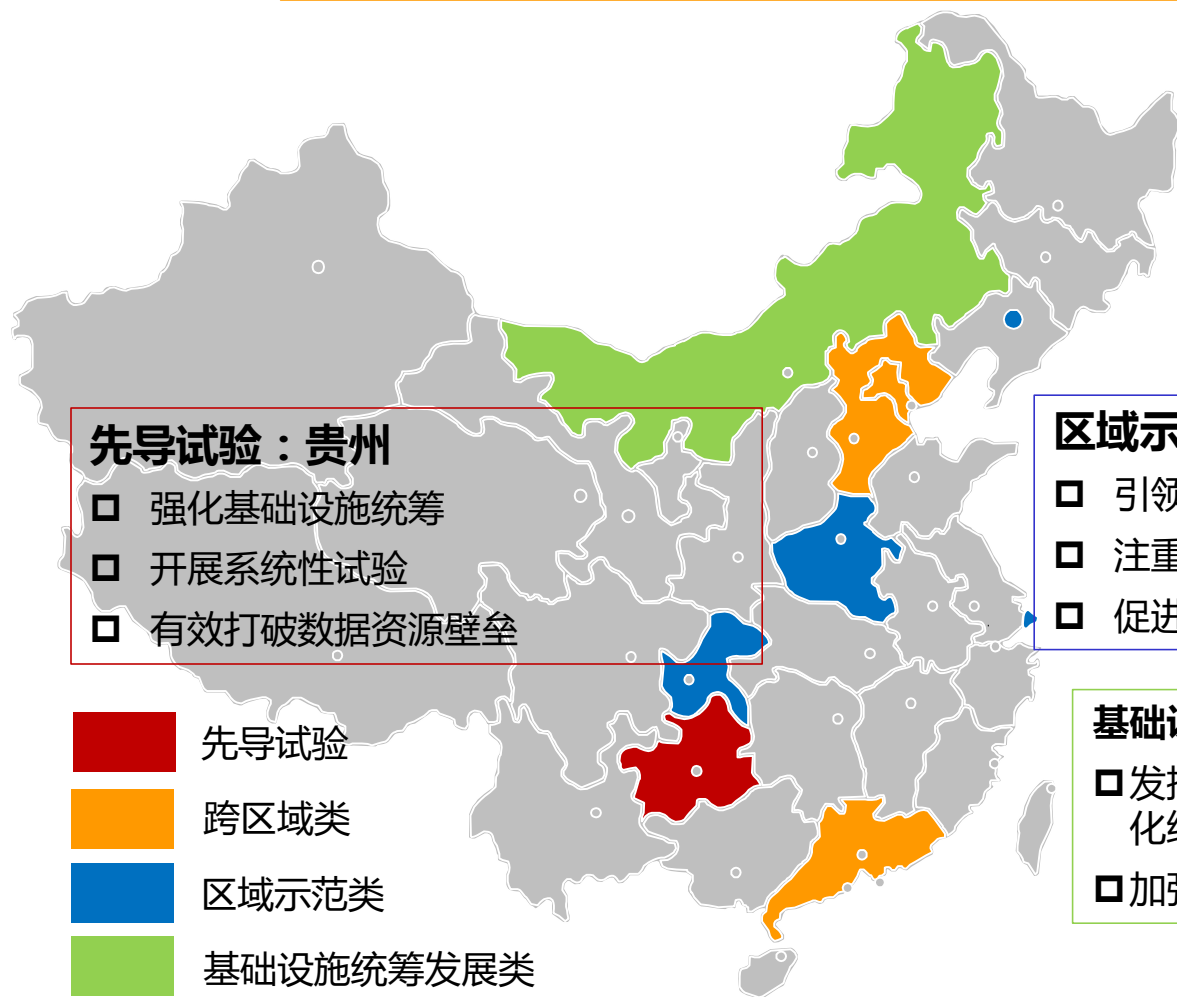
## 试验目标

- ❑ 通过3~5年时间探索，推动制度、管理、服务和协同创新，有效打破数据资源壁垒，形成若干辐射带动效应强、示范引领作用显著的区域性试验基地
- ❑ 打造一批大数据先进产品，培育一批大数据骨干企业，建设一批大数据众创空间，培养一批大数据产业人才，有效推动相关制度创新和技术创新，发掘数据资源价值，提升政府治理能力，推动经济转型升级

# 全国多片区布局综合试验区

## 批复 历程

- 2016年2月25日，三部委批复**贵州**建设国家大数据(贵州)综合试验区。
- 2016年10月8日，三部委批准**7个区域**建设国家大数据综合试验区。



# 全国各地方纷纷制订各自大数据发展政策规划

据中国信息通信研究院统计，截止到2017年1月底，共有16个省、21个市发布了大数据发展规划

省市	文件名称	省市	文件名称	省市	文件名称
北京	北京市大数据和云计算发展行动计划	广州	关于促进大数据发展的实施意见	沈阳	沈阳市促进大数据发展三年行动计划
上海	上海市大数据发展实施意见	深圳	深圳促进大数据发展行动计划	哈尔滨	促进大数据发展若干政策（试行）
江苏	江苏省大数据发展行动计划	青岛	关于促进大数据发展的指导意见	合肥	合肥市大数据发展行动纲要
浙江	浙江省促进大数据发展实施计划	宁波	关于推进大数据发展的实施意见	淮南	大数据产业发展三年行动计划
广东	广东省促进大数据发展行动计划	南京	南京促进大数据发展三年行动计划	东莞	东莞市大数据发展实施方案
山东	关于促进大数据发展的意见	武汉	武汉大数据产业发展行动计划	兰州	促进大数据发展的实施意见
安徽	“十三五”软件和大数据产业规划	郑州	促进大数据发展行动计划	赣州	加快大数据发展实施意见
广西	促进大数据发展的行动方案	南宁	南宁市大数据建设发展规划	盐城	推进大数据产业发展的实施意见



# 全国各地确立符合发展情况的大数据发展定位

## 积极面向全国需求推进大数据应用与产业创新

	面向全国	面向地方
产业中心/高地	北京、重庆、 <b>广东</b> 、浙江、湖北、新疆、深圳、武汉、沈阳、宁波、天津滨海新区、贵安新区	青岛、合肥、淮南、台州
应用示范中心	北京、贵州、 <b>广东</b> 、浙江、湖北、深圳、青岛、宁波、淮南、贵安新区	青海、台州
资源聚集中心	贵州、湖北、青海、武汉、天津滨海新区	沈阳、兰州、淮南
创业创新中心	北京、 <b>广东</b> 、青岛、宁波、贵安新区	淮南
人才发展中心	湖北	

- 在发布大数据规划的37个省市中，明确定位的有20个，占统计总量的54%
- 12个省市提出建设面向全国的大数据产业中心或高地，产业发展受关注度最高
- 东部、中部、西部不同区域的省市均把大数据作为新的经济与产业增长点



# 全国各地区域发展特色与优势

## 五大区域构成我国大数据产业总视图

### 中西部地区

- 借助环境、能源、价格优势，实现集群效应，打造经济发展新动能
- 建设数据中心、网络基础设施与大数据产业园区

### 珠三角地区

- 以广州、深圳为引领，打造全国大数据应用先导区，建设全国的大数据产业聚集区，探索大数据的新路径、新经验

八大综合试验区，引领大数据发展新思维

内蒙古：大数据基础设施统筹发展综试区

京津冀：跨区域综试区

沈阳

河南

上海

重庆

贵州：首个综试区

珠三角：跨区域综试区

### 东北地区

- 以大数据发展为主体，以传统产业转型升级和智慧城市建设为两翼，完善大数据产业链，振兴东北老工业基地

### 环渤海地区

- 发挥北京的人才与科技优势，天津、河北、山东的制造业与资源优势，建设国家大数据产业创新中心、全球大数据产业创新高地

### 长三角地区

- 依托上海、江苏、浙江、安徽等地在金融、ICT、互联网、制造业等领域优势，实现大数据与智慧城市、云计算协同发展

# 广东省发布促进大数据发展行动计划

**总目标：用5年左右时间，打造全国数据应用先导区和大数据创业创新集聚区，抢占数据产业发展高地，建成具有国际竞争力的国家大数据综合试验区。**

**大数据基础设施建设、资源整合和政府数据开放共享取得显著成效。**

建成全省统一的电子政务数据中心，以及10个左右地市级政务数据分中心。

**大数据创新应用深入经济社会各领域。**

运用大数据推动形成精准治理、多方协作的社会治理新模式

**大数据产业成为重要的经济增长极。**

基本形成高端智能、新兴繁荣的大数据产业发展新生态和大众创业、万众创新的创新驱动新格局

---

## 二、大数据产业生态与应用实践案例

- 大数据产业生态分析
- 大数据技术发展分析
- 大数据在经济生活各领域的应用

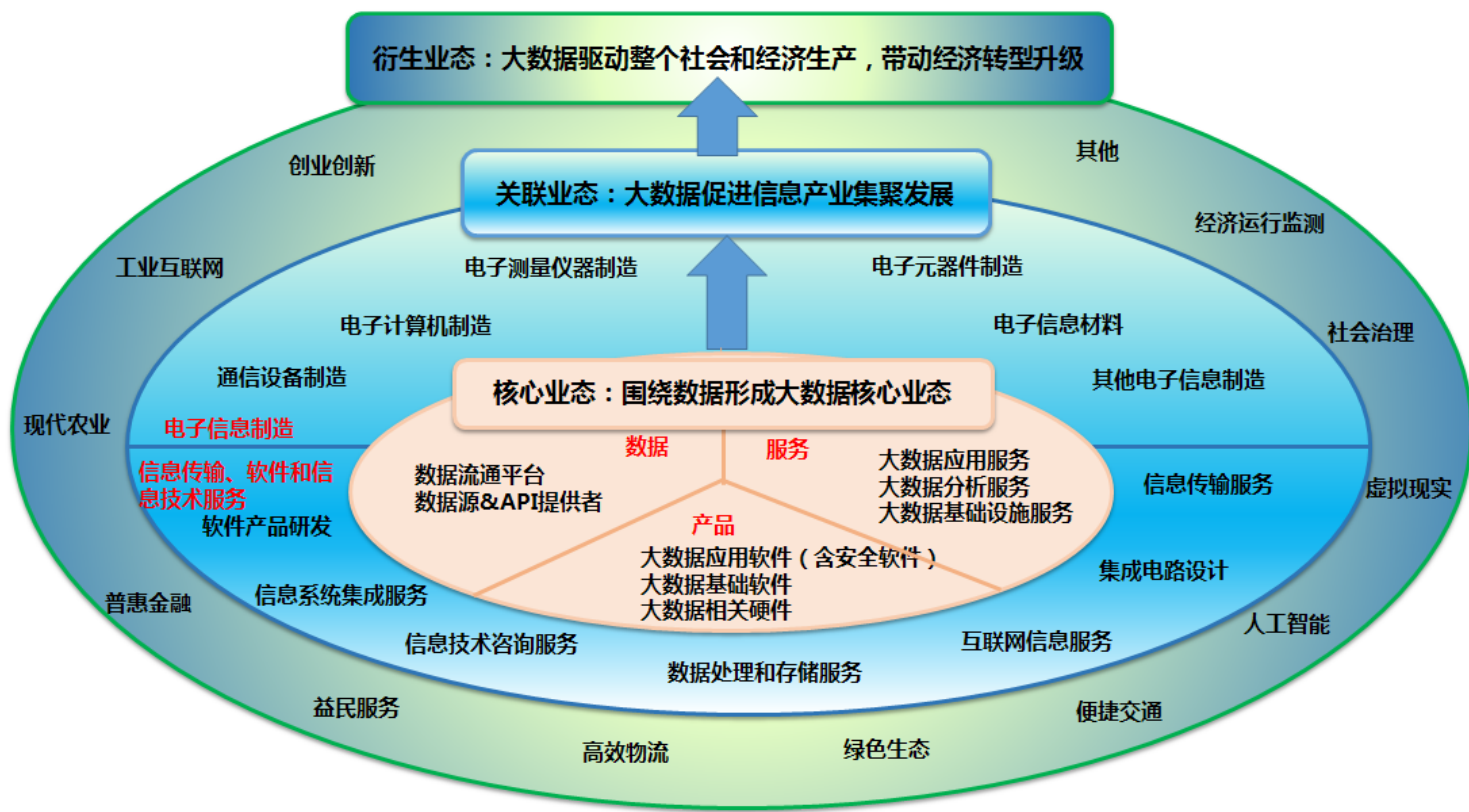
---

## ●大数据产业生态分析

# 《大数据产业发展规划》对大数据产业界定

大数据产业由大数据核心业态、大数据关联业态和大数据衍生业态构成

- 核心业态，包含数据资源提供、大数据管理和分析软硬件及服务等环节
- 关联业态，对大数据提供基础支撑的电子信息产业，包括电子信息制造，信息传输、软件和信息技术服务等
- 衍生业态，通过大数据技术与传统产业的应用和融合而衍生出新型业态

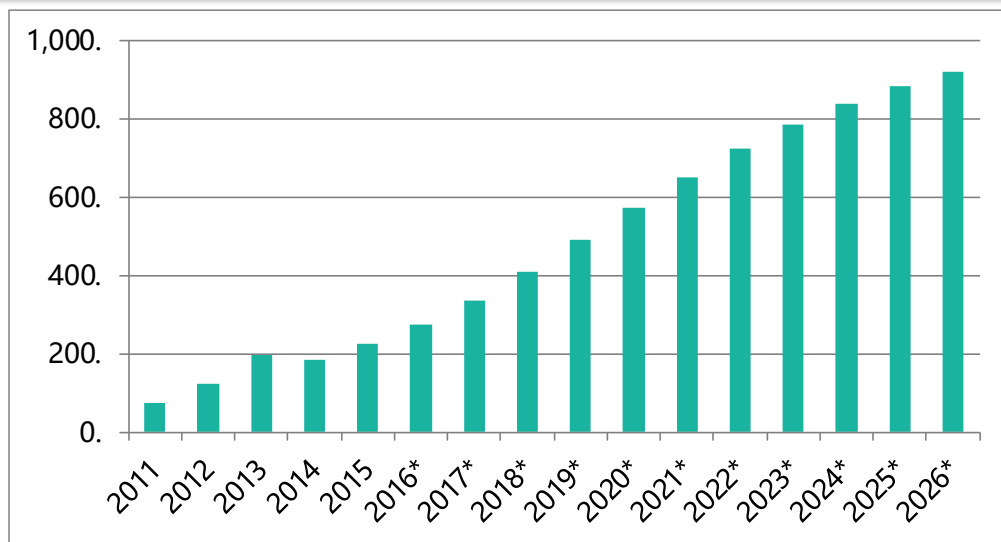


# 大数据核心产业构成

**大数据核心业态**包括“**上游数据+中游产品+下游服务**”，产业链上下游各类参与者共同构成大数据核心产业体系。



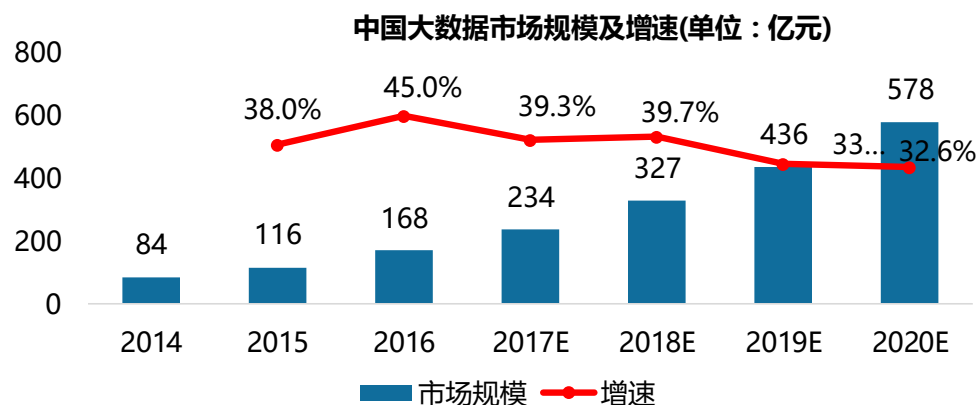
# 全球及我国大数据市场规模（核心产业）



数据来源：Wikibon，2016年3月，单位：亿美元

根据 IDC、Wikibon 等咨询机构分析，2016年，全球的大数据核心产业规模约为300亿美元，2000亿人民币左右。

根据信通院的市场调查，2016年中国大数据市场规模约为168亿人民币。



数据来源：大数据市场调查报告（中国信通院，2017.03）

# 我国大数据企业区域分布



全国大数据企业热力分布图  
(企业数量+叠加注册资本)

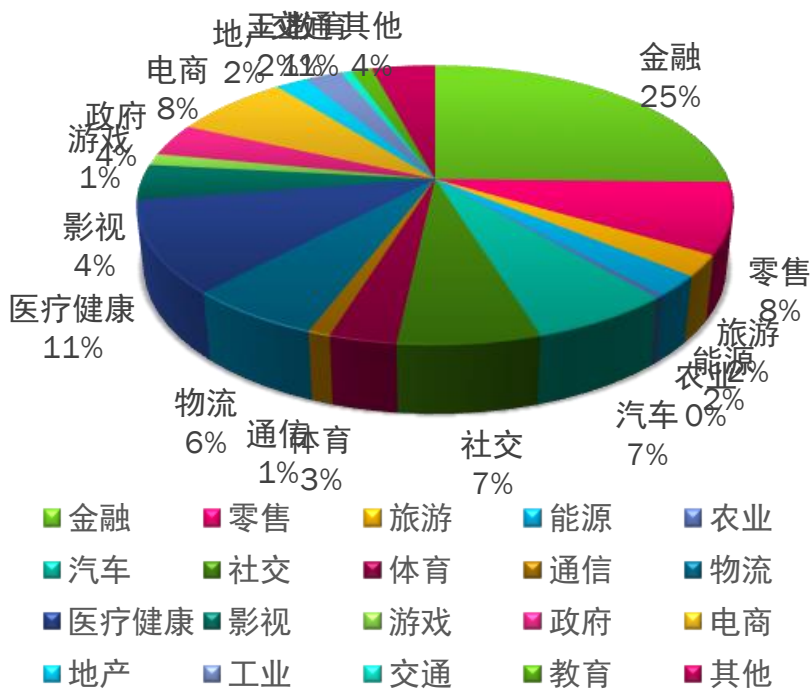
数据来源：中国大数据和人工智能产业分析平台（中国信通院，2017.03）

如图所示：北京地区、长三角、珠三角及成渝经济圈是大数据企业热度最高的区域。

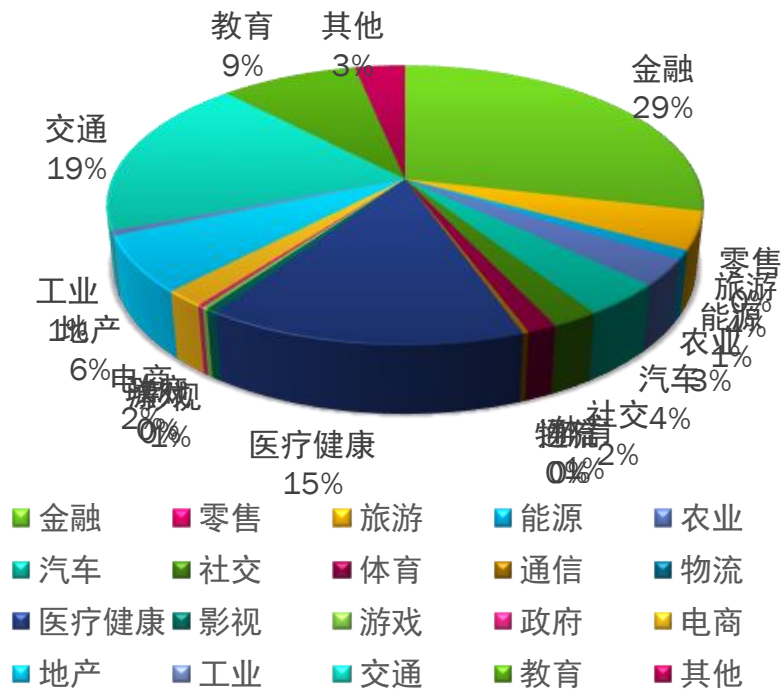


# 大数据应用领域渗透分析

## 行业服务渗透分析



## 行业应用渗透分析



排名前三的领域：金融、医疗、社交

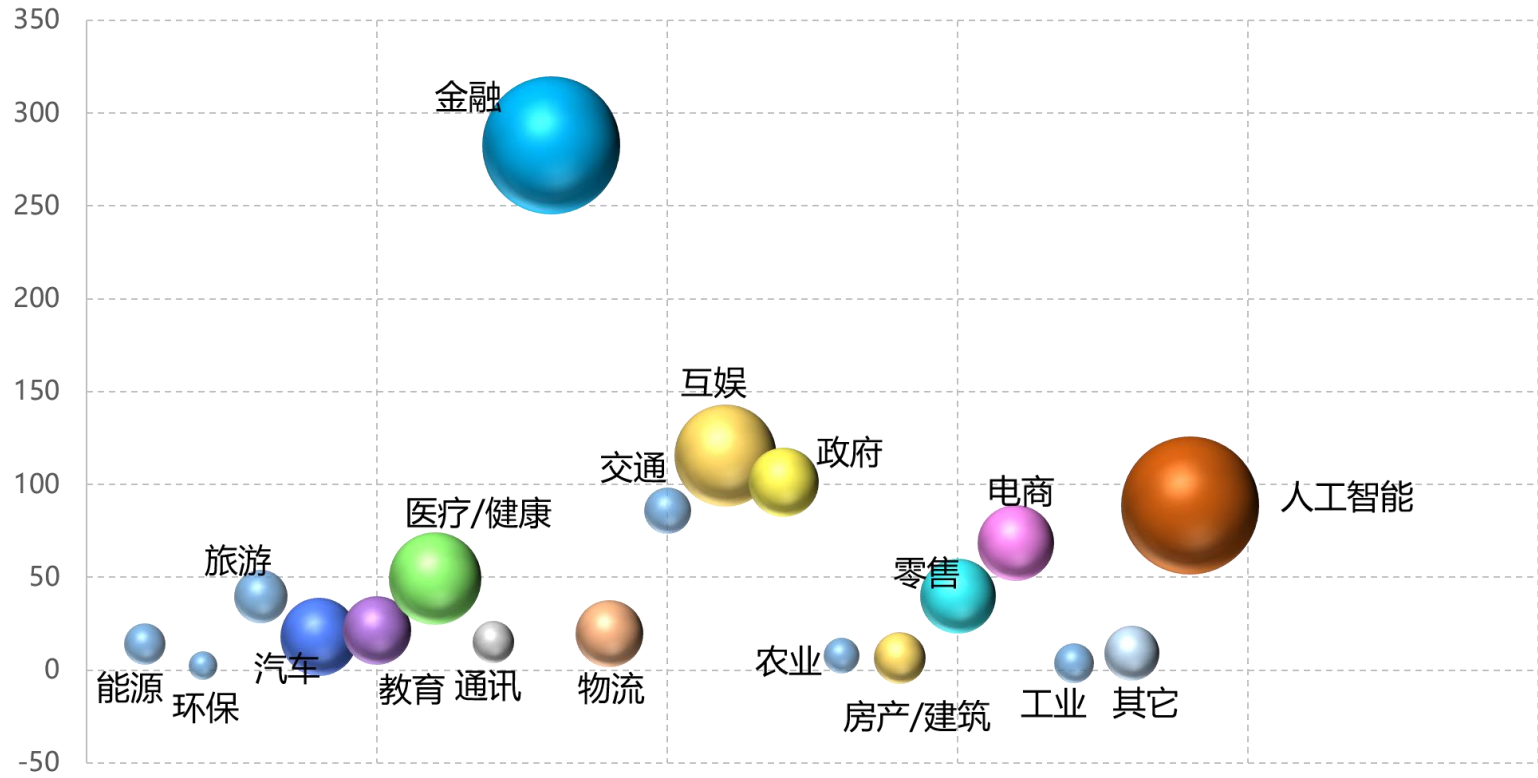
排名前三的领域：金融、医疗、交通

数据来源：中国大数据和人工智能产业分析平台（中国信通院，2017.03）

# 大数据金融行业融资火爆，人工智能蓄势待发

2012-2016 大数据一级行业融资分布

单位：亿元



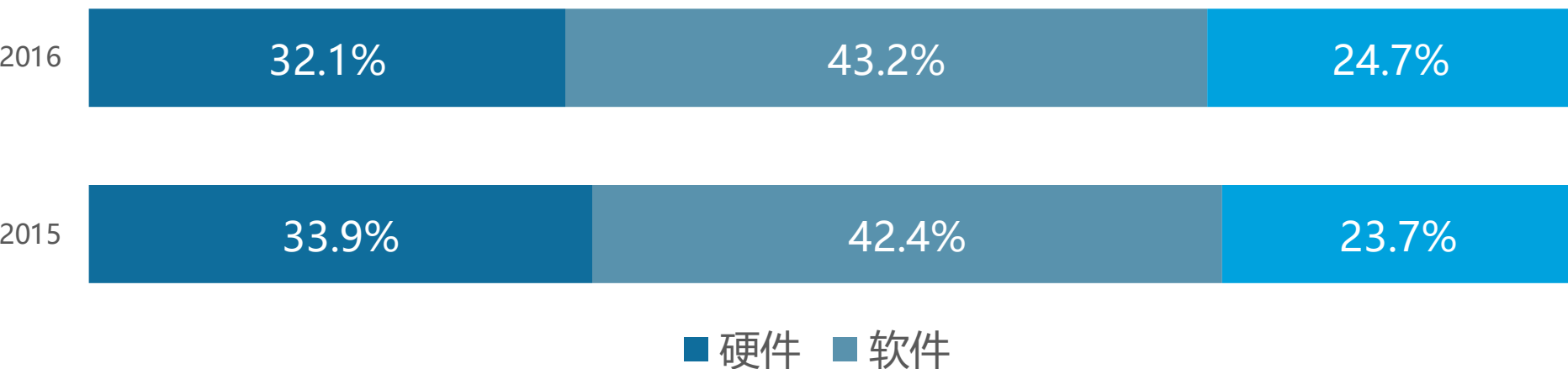
数据来源：中国大数据和人工智能产业分析平台（中国信通院、数据猿，2017.03）

2012-2016五年间，金融行业 融资事件为221起，已披露融资金额为283亿，在19个一级行业中遥遥领先；  
人工智能 行业融资项目数与金融行业持平。

# 大数据细分市场逐年变化

根据中国信息通信研究院的统计，2016年与2015年相比，硬件市场规模为53.9 亿元，占比下降1.8%；软件市场规模为72.6亿元，占比提高0.8%；大数据服务市场规模41.5亿元，占比提高1.0%。这体现出大数据市场细分领域中，软件和服务比重呈上升趋势，硬件比重逐年减少。

中国大数据细分领域构成



# 广东省大数据产业发展情况（1）

目前广东省大数据产业分布呈现“广深引领发展、珠三角加快发展、粤东西北起步发展”的态势。

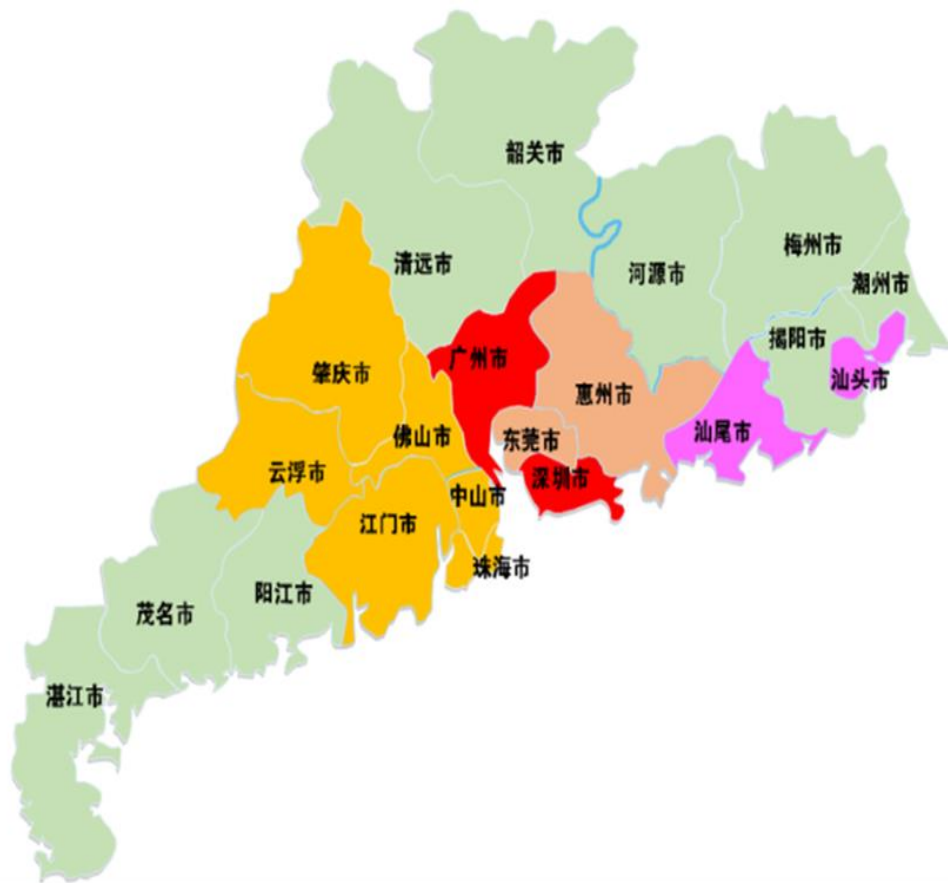
广州、深圳率先布局大数据产业发展规划建设，2015年广深两市大数据产业产值占全省近90%，其中广州39.6%，深圳50.3%。

目前，广州已拥有亚太信息引擎、中国电信沙溪云计算中心等；深圳已建设有华为科技城新一代通信产业基地、深圳软件产业基地、国家电子信息产业基地、国家电子商务示范城市等基地。

广东省其他地区也在依托本地优势加快大数据产业的发展，积极开展大数据产业布局，如佛山的云计算中心，东莞的云存储中心、肇庆的云服务产业园，江门的“珠西数谷”项目，云浮市的云计算数据中心产业园（“云谷”）、韶关市的“韶关市翔计划”、深汕特别合作区大数据产业园、清远市与阿里巴巴、腾讯、广东铁塔等公司达成的战略合作。



# 广东省大数据产业发展情况（2）



2016年10月，广东省获批珠三角国家大数据综合试验区，将重点打造“一区两核三带”的总体功能布局，全面推动全省大数据的发展。“一区”指珠三角国家大数据综合试验区，“两核”指广州、深圳大数据产业核心集聚区，“三带”分别是珠江西岸智能制造大数据产业带（佛山、珠海、中山、肇庆、江门、云浮）、珠江东岸大数据技术和产品制造产业带（东莞、惠州）、沿海及跨境大数据产业带（汕头、汕尾）。

---

## ●大数据技术发展分析

# 大数据技术快速演进，底层技术日趋成熟

阶段1：RDBMS一枝独秀  
One size fits all  
SQL/ACID

阶段2：NoSQL百花齐放  
One Size Does Not Fit All  
NoSQL/CAP

阶段3：回归SQL  
One Size Does Not Fit All  
SQL/NewSQL

面向企业应用的商业智能提出，面向数据仓库的数据分析（OLAP）技术兴起

开源Apache Hadoop逐渐兴起，大幅推动互联网大数据应用

Spark、Flink等新一代融合大数据平台，SQL on Hadoop兴起

1995

2000

2005

2010

2015

分布式技术提出  
谷歌提出分布式文件系统、分布式数据库和分布式计算框架，奠定大数据技术基础

实时计算技术提出  
流计算、图计算、交互式分析、内存计算等技术不断演进

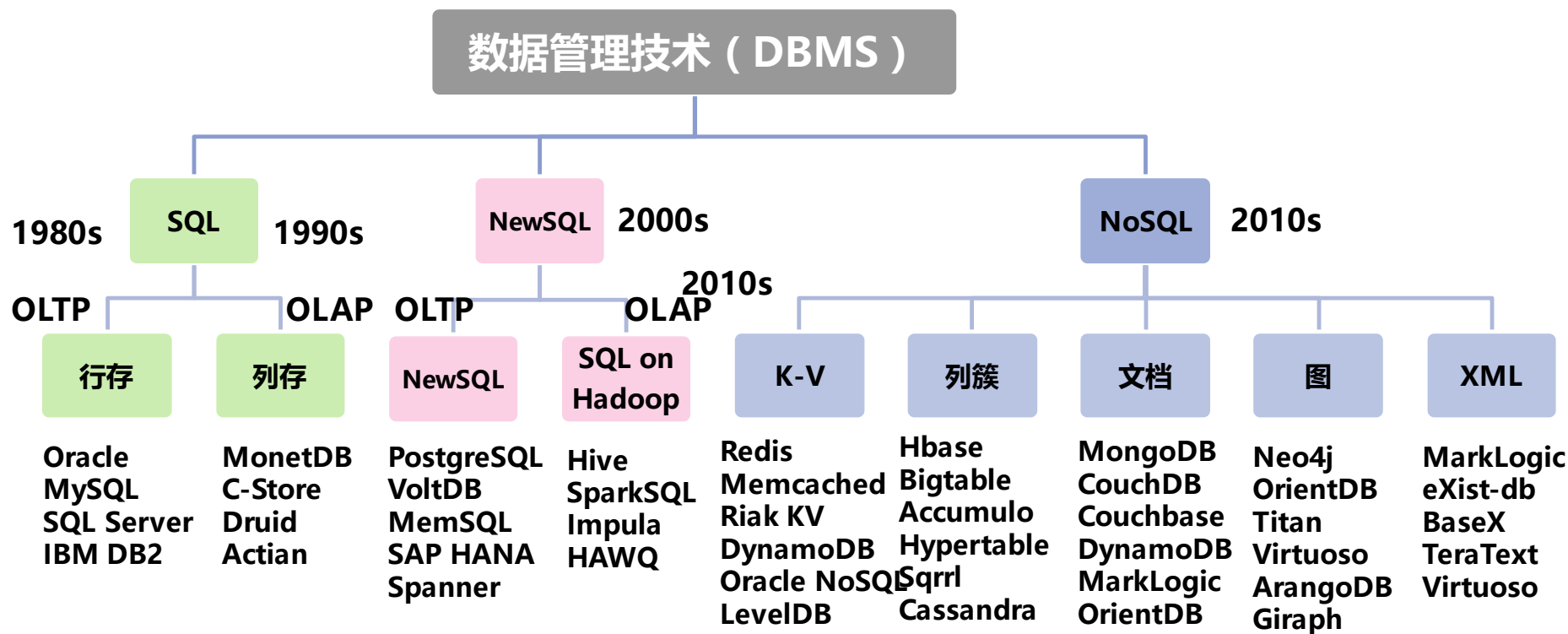
前大数据时代

底层分布式技术日趋成熟，释放上层数据分析技术的创新活力：  
实时流分析、交互分析、机器学习



# 数据平台技术创新活跃，解决方案百花齐放

大数据需求拉动技术创新，衍生出310多种不同的数据管理技术



**对上采用统一接口，底层使用内存技术成为主要趋势：**

1. SQL再次流行，大数据平台向数据管理平台发展，SQL on Hadoop技术成为
2. SSD和Flash广泛使用，颠覆了之前数据库设计的假设，内存数据库发展迅猛，如HAHA，Spark
3. 交易（OLTP）型和分析（OLAP）型相互融合，产生新的融合数据管理平台技术（HTAP）
4. 各种NoSQL技术百花齐放，one size dose not fit all



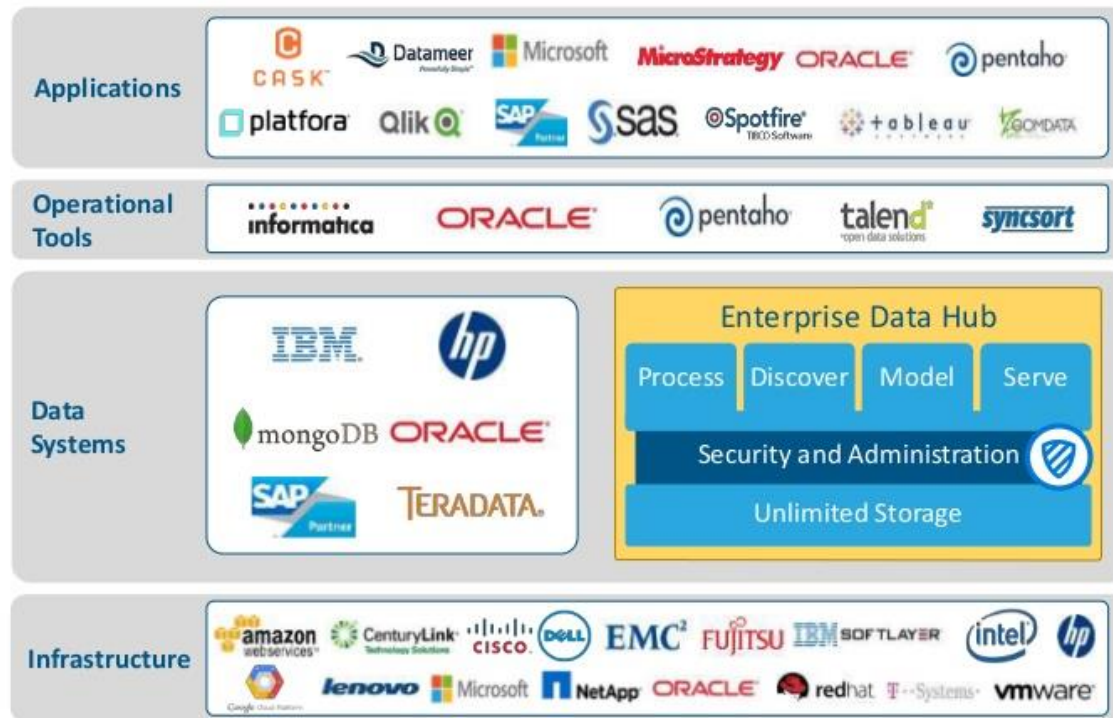
# 大数据平台产品竞争激烈，进入拼生态阶段

未来只有少数平台存活，竞争日益激烈

Hadoop平台产品前四强：Cloudera、Hortonworks、MapR和IBM

如：Cloudera不断扩大兼容性列表，Hortonworks倡导Hadoop的标准化

None of this is possible without the ecosystem



More than 1,600 partners ensure compatibility with existing investments, lower skill barriers, and help maximize value from your data.

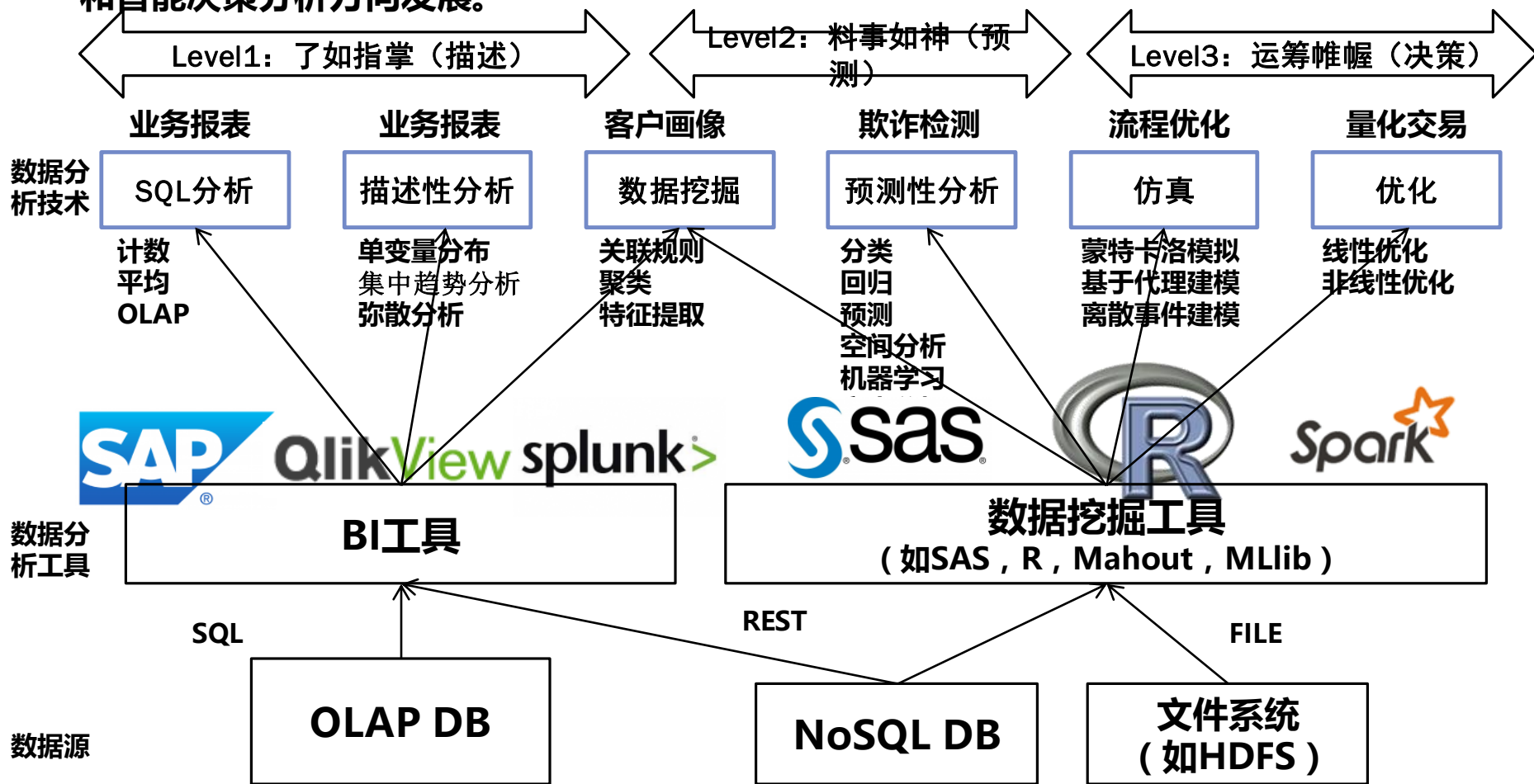


cloudera

© Cloudera, Inc. All rights reserved. 16

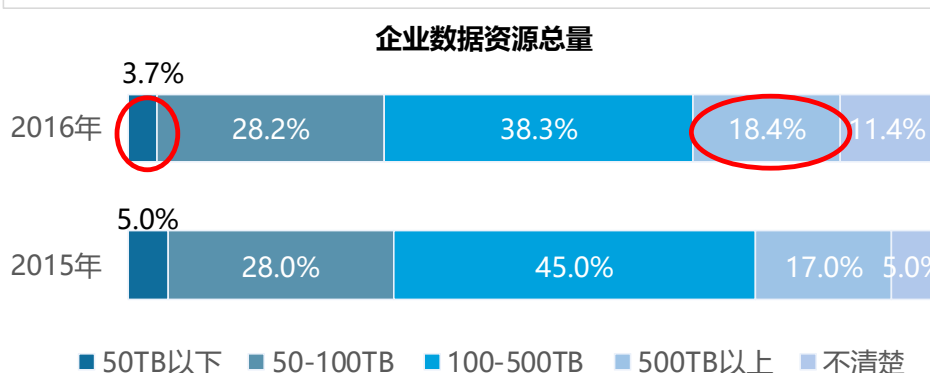
# 大数据分析技术：从了如指掌到运筹帷幄

大数据分析技术：正在从基于统计方法的描述性分析，向基于机器学习的智能预测分析和智能决策分析方向发展。

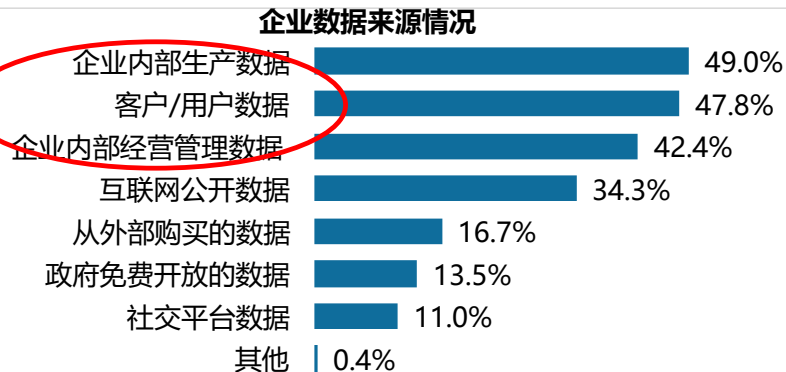


# 大数据的资源现状

## 企业数据资源规模增长迅速，企业数据资源的主要来自于企业内部

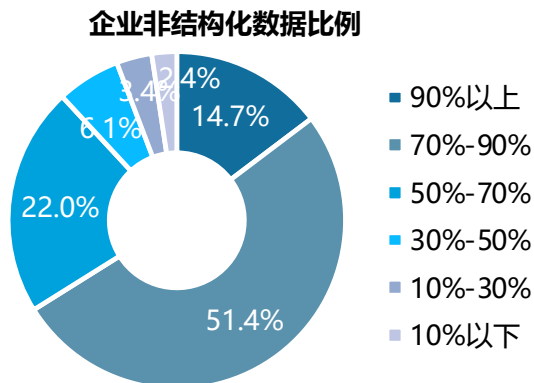


数据来源：中国信息通信研究院 2017,03

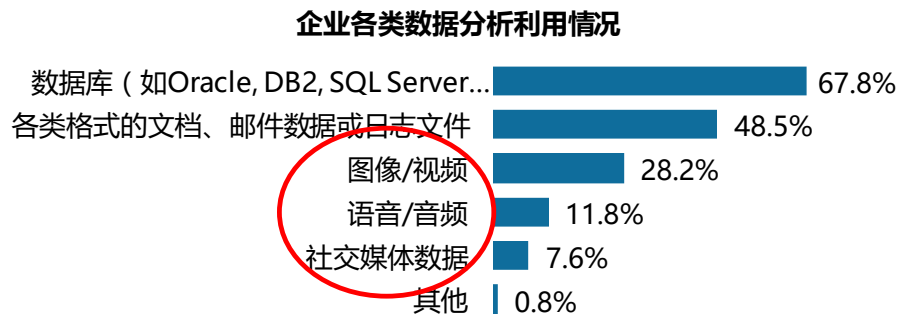


数据来源：中国信息通信研究院 2017,03

## 绝大部分数据资源为非结构化数据，而图像/视频、语音/音频、社交媒体数据等非结构化数据的分析和利用有待提高



数据来源：中国信息通信研究院 2017,03



数据来源：中国信息通信研究院 2017,03

---

- **大数据在经济生活各领域的应用**

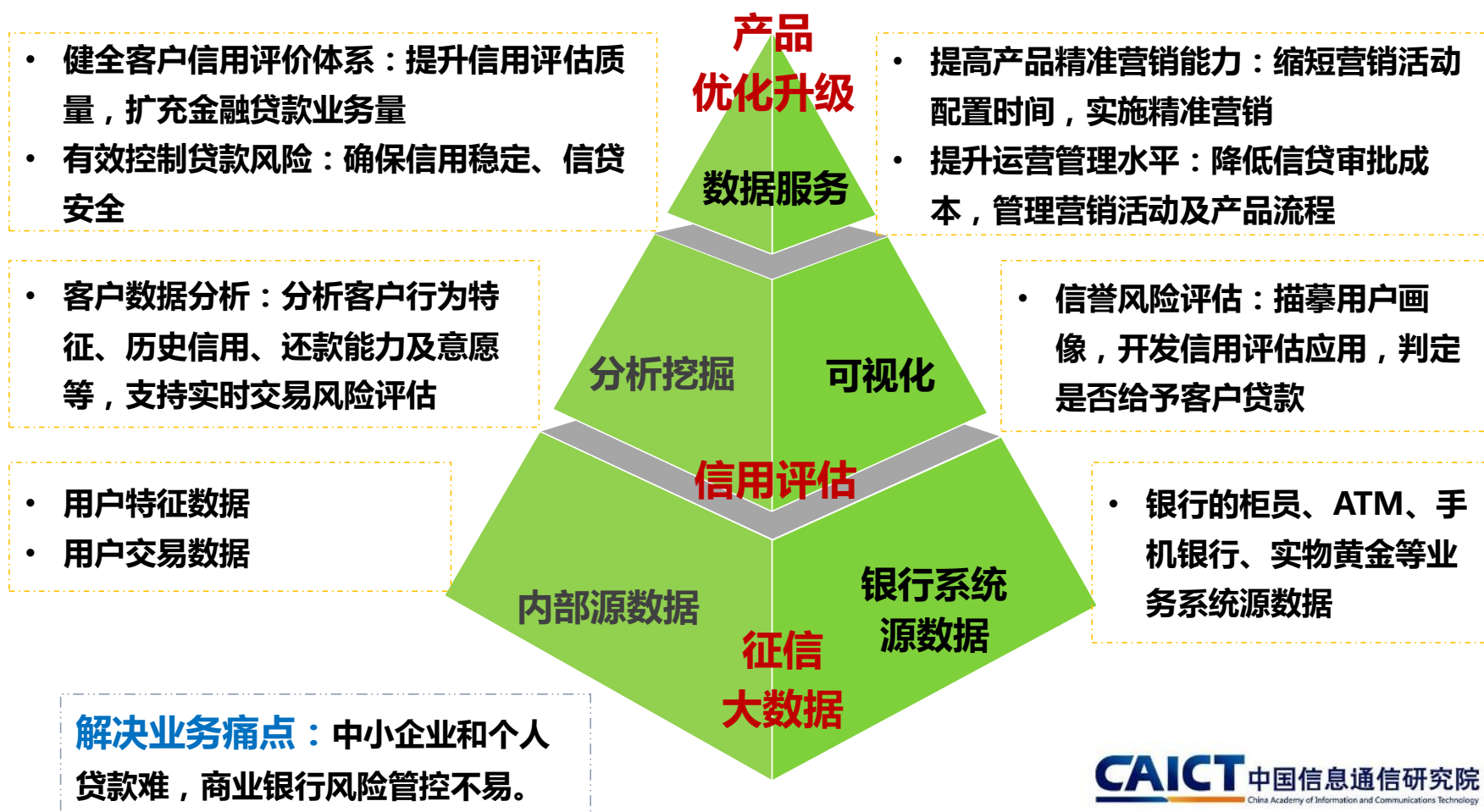
# 精准营销大数据

营销行业是较先应用大数据的行业，已形成成熟的变现模式。



# 金融风控大数据

金融、征信、风控领域也是应用大数据非常成功的领域。综合分析、挖掘电信和金融双方数据，及时把握客户行为并进行信用评估。





# 工业大数据

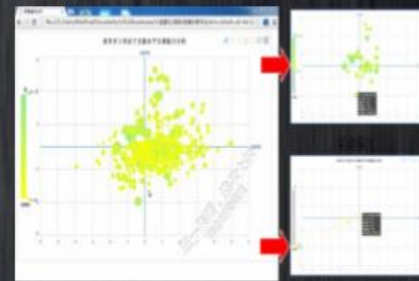
工业大数据是我国加快实现制造强国的重要支撑力。



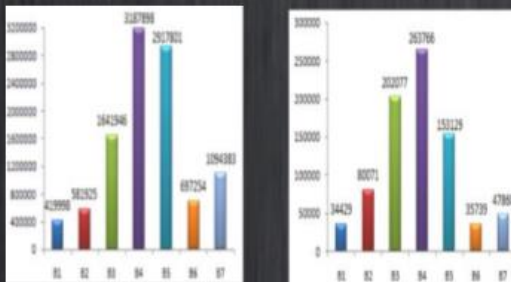
宏观经济预测



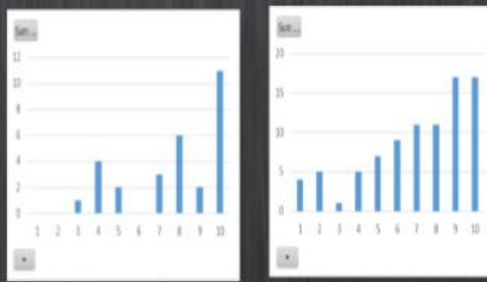
配件需求预测



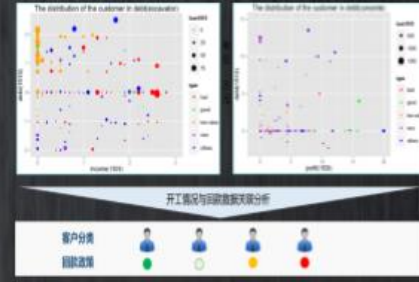
商业（服务）模式创新



指导产品研发



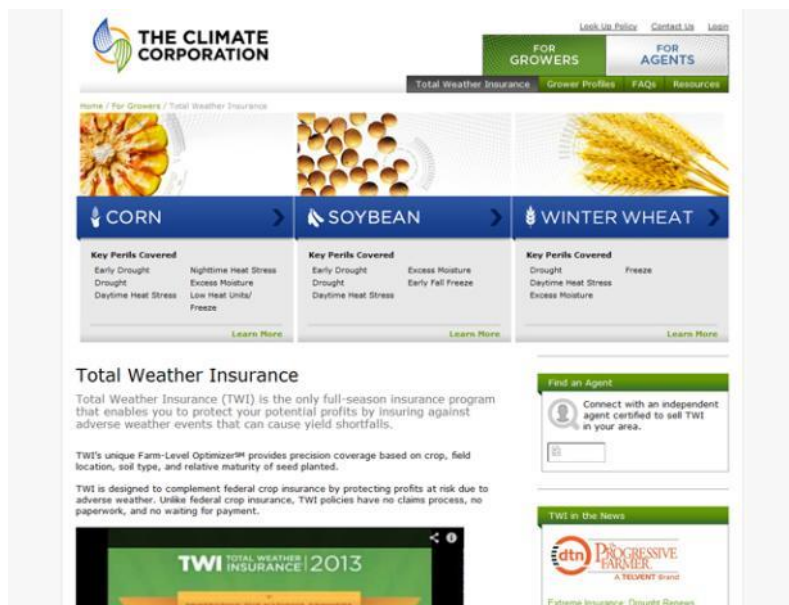
设备故障预测



金融风险管控

# 农业大数据

大数据指导农业生产可提高农业生产科学性，有效降低农民风险。



- 全气候保险 ( TWI )
- 雹灾保险 ( Crop-Hail )
- 干旱保险 ( Named Peril Drought )
- 大数据平台 ( Climate Basic|Pro )

大数据技术

气候学

农业  
经济学

2500万  
气候度量指标

1500亿次  
土壤监测数据

日数据  
气候模型

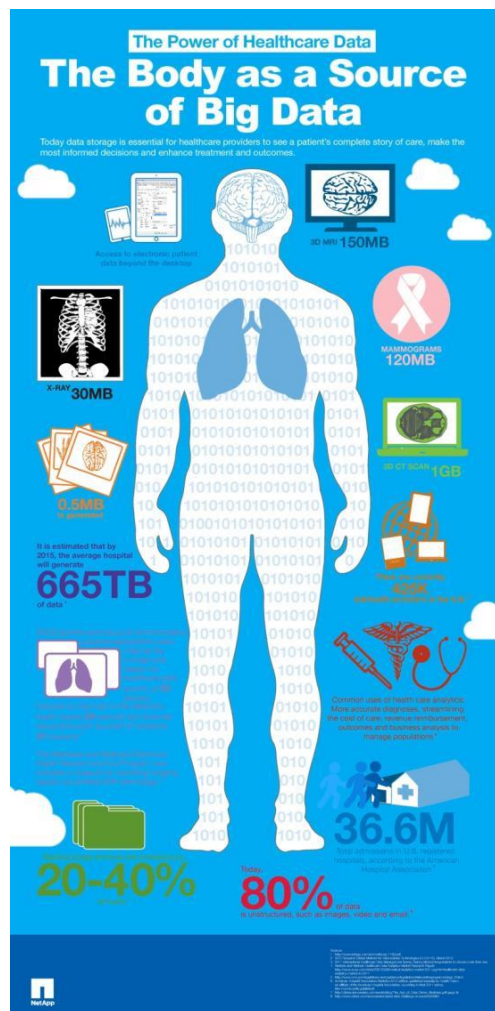
10万亿  
综合气候数据

50TB  
大数据处理能力



# 医疗健康大数据

大数据在医疗健康领域应用极广。



业务应用	内容
临床数据比对	匹配同类型的病人的用药情况，分析最佳治疗途径；
临床决策支持	（药物过敏、重点人群、慢病患者等各类警示信息以及重复检验/检查提示等）
实时统计分析	展示公共卫生统计数据
远程病人数据分析	监控，分析临床监护数据
人口统计学分析	对不同体质人群分类
就诊行为分析	跟踪健康卡数据，分析病人就诊行为
基本药物临床应用分析	分析基本药物在处方中的比例
药品研发	基于大数据的药品市场预测
新农合基金数据分析	辅助制定农合基金的起付线，赔付病种等
新的服务模式	根据临床数据分析，提供个性化服务及自助服务等新模式

# 城市规划大数据

抓取、分析、挖掘内部及各类互联网监控数据，掌握目标地区及周边用户特征，准确、高效地实现预期利益最大化的商业选址及服务内容决策。



## 周边用户特征数据

- 周边用户特征分类数据
- 人流量、交通运行数据
- .....



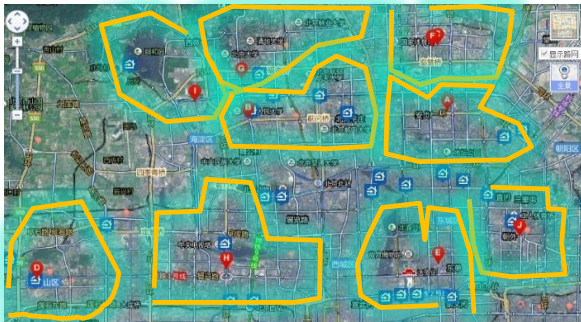
## 分析挖掘与可视化

- 基于GIS技术：城市区域栅格化处理，分析每个栅格特征数据，挖掘店铺选址决策依据；
- 结合可视化技术：实时展示店铺周边人群关注内容等情况。



## 支撑商业选址决策

- 选址分析与推荐：选址分析、优劣对比及最优化推荐；
- 其他数据服务：可提供商圈分类、位置管理、客户挖掘、竞品分析等支持。



# 交通大数据

重点利用位置数据，提供交通拥堵实施监控、人流实施监控等辅助管理服务。

## 交通实时监控

客户需求：通过信令数据为广州交委提供广州火车站、火车东站等特定目标区域内的实时人流量统计、辅助进行应急指挥。

### 热点区域交通拥堵分析



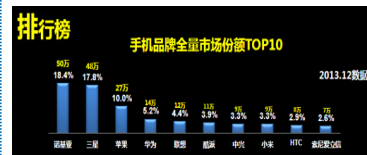
## 数据报告

- 《人口流动研究报告》
- 《人口分布研究报告》
- 《公共交通研究报告》
- 《人口特征研究报告》
- .....

## 人流实时分析

客户需求：深圳市发改委区域经济处牵头组建“全口径人口信息集成应用系统”，基于信令数据，计算全市人口数量、分布、流动等。

### XX社区人口变化曲线图



# 公安大数据

大数据技术正在深刻改变公安在信息收集、反恐维稳、打击犯罪等方面的能力。

## 加强情报研判应用

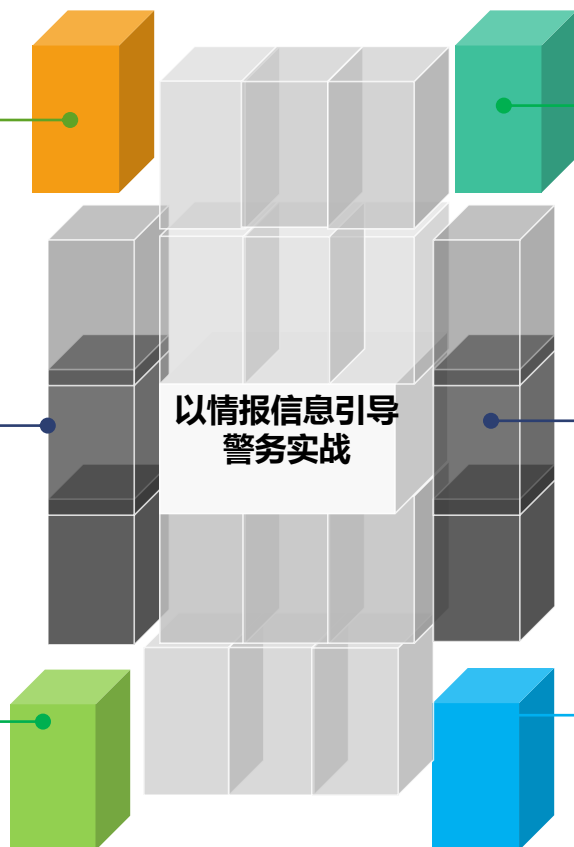
- 纵向贯通、横向集成、高度共享、整体联动

## 提升反恐维稳水平

- 应用大数据技术，构建反恐情报研判模型，形成主动维稳预警格局

## 提高打击犯罪能力

- 利用大数据警务云，打好信息战、合成战，精确打击、提升打击犯罪的工作的能效。



## 构建互联网络综合防控体系

- 完善舆情监控体系，加强网上网下数据融合、分析、挖掘能力

## 推进执法与监督信息化

- 拓展网上执法办案，深化网上办案执法效果，建设网上督查系统

## 开展信息便民惠民服务

- 网上便民服务平台，实现网上全流程办理，数据多跑路，群众少跑腿

# 应用案例：阿里巴巴用大数据进行用户画像

**应用场景：基于网络行为大数据的用户画像，个性化精准化服务**

(1) 采集用户电商、浏览、社交等数据 → (2) 分析用户特征喜好，给用户贴上千个标签

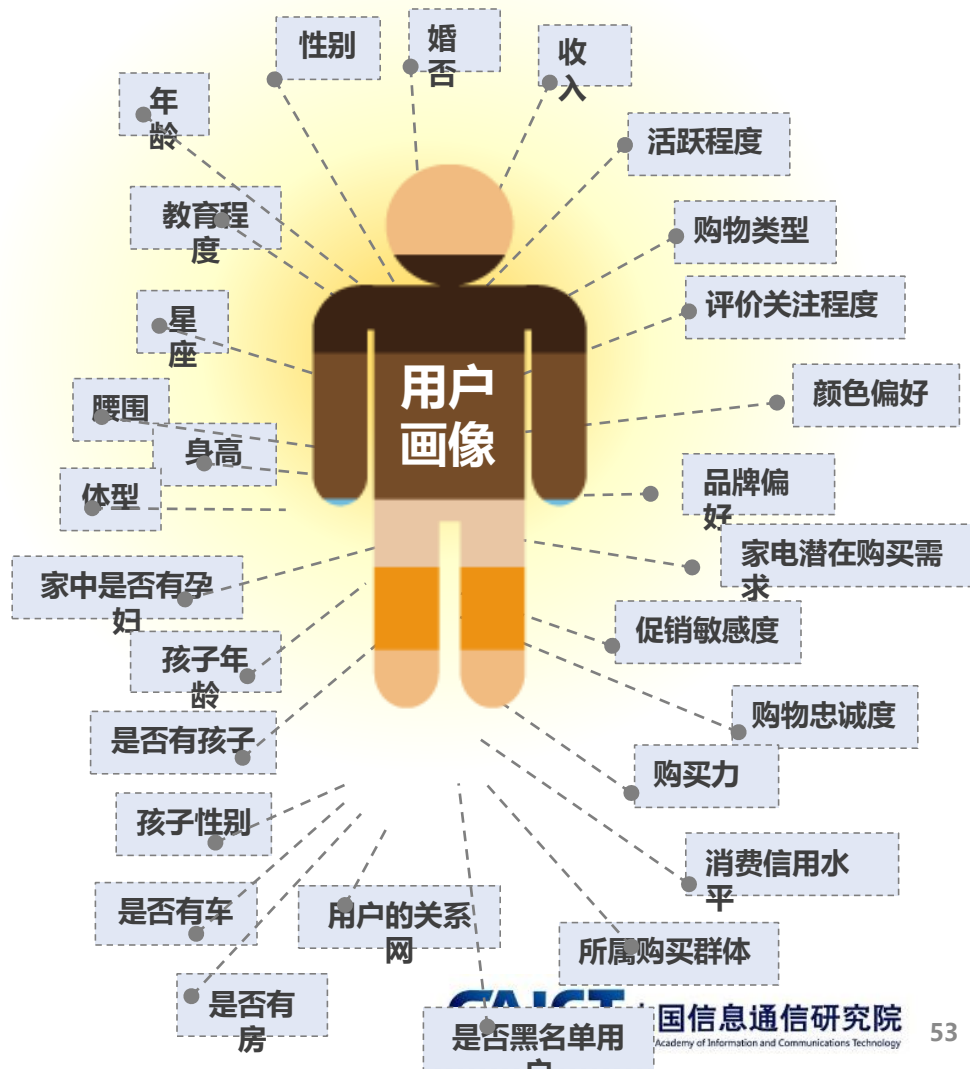


阿里、腾讯等的数据量：100PB量级

(3) 应用：个性广告、精准营销、**信用评估**、反欺诈等



基于数据的  
芝麻信用





# 应用案例：大数据促进“低库存经济”诞生

**“低库存经济” 概念产生于C2B模式**

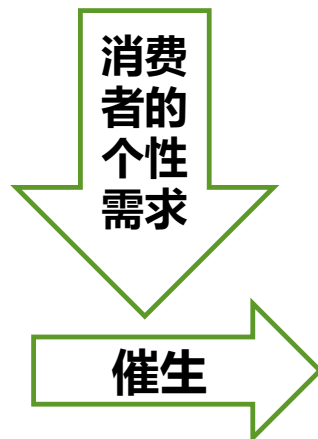


2016年1月27日，李克强总理在的国务院常务会议上指出，“所谓的‘C2B’，就是消费者提出要求，制造者据此设计消费品、装备品。这是一场真正的革命：一个企业不再是单个封闭的企业了，它通过互联网和市场紧密衔接，和消费者随时灵活沟通。”

**低库存经济即在相对较低的库存条件下实现生产过程平稳运行和经济效益不断提升的经济过程。**



**大批量生产降低边际成本  
传统经济**



**大规模实现个性化定制的模式才是“低库存经济”的精髓。**

---

## 三、大数据推动政府治理变革

- 大数据推动的社会变革与政府治理变革
- 政府大数据建设的推进步骤与目标

Why  
How

---

- **大数据推动的社会变革与政府治理变革**



# 时代背景：信息革命与数字化浪潮



从社会发展史看，人类经历了农业革命、工业革命，正在经历信息革命。

信息革命则增强了人类脑力，带来生产力又一次质的飞跃，对国际政治、经济、文化、社会、生态、军事等领域发展产生了深刻影响。

——习近平总书记2016年4月19日在全国网络安全和信息化工作座谈会的重要讲话

当今世界，网络信息技术日新月异，全面融入社会生产生活，深刻改变着全球经济格局、利益格局、安全格局。世界主要国家都把互联网作为经济发展、技术创新的重点，把互联网作为谋求竞争新优势的战略方向。

要加快传统产业数字化、智能化，做大做强数字经济，拓展经济发展新空间。

——习近平总书记2016年10月9日在中共中央政治局第36次集体学习的重要讲话

# 时代背景：国家部署“互联网+”行动

## 内涵 要义

“互联网+”就是把互联网的创新成果与经济社会各领域深度融合，推动技术进步、效率提升、组织变革，提升实体经济的创新力和生产力，形成更广泛的以互联网为基础设施和创新要素的经济社会发展新形态。

## 近期 目标

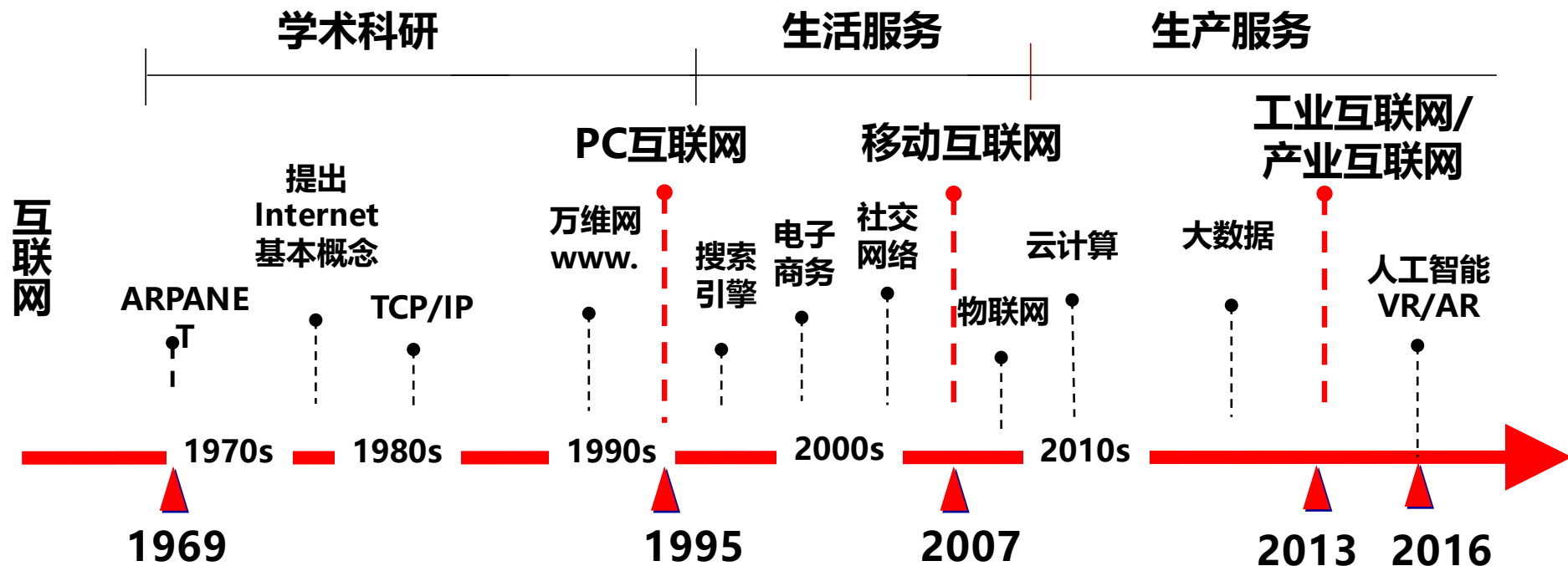
2018年，互联网与经济社会各领域的融合发展进一步深化

## 远期 目标

2025年，“互联网+”产业生态体系基本完善，“互联网+”新经济形态初步形成，“互联网+”成为经济社会创新发展的重要驱动力量

# 以互联网为代表的信息技术成为重要的变革力量

互联网应用：消费互联网 → 产业互联网/互联网+



互联网改变了世界，还改变了世界改变的方式.....

千年的文化传统几年内就能够改变外在形态：**拜年、发红包**.....

根据腾讯发布的数据显示，除夕当日全国共收发微信红包142亿个，比去年增长75.7%；  
零点前后，微信红包祝福每秒76万个

# 生产要素变革：数据和信息成为新生产要素

以数据、知识、技术、信息、等新生产要素为支撑的经济发展新动能正在形成——《培育壮大经济发展新动能加快新旧动能接续转换的意见》

	农业时代	工业时代	信息时代
生产要素	劳动力	资本、劳动力	信息、资本、劳动力
基础设施	土地	铁路、公路、电网	互联网
主导产业	农业	工业	信息产业

数据已成为除劳动力、资本之外的新的生产要素，广泛渗透到经济社会各领域

互联网已成为经济社会发展的战略性基础设施

以数据产业为代表的信息产业成为20世纪90年代以来，推动全球经济增长、财富创造和就业增加的主导性产业，是“新经济”的核心力量。

# 创新模式变革：依托大数据和互联网平台加快迭代

传统创新模式	融合创新模式	大数据线上创新模式
<div>Metersbonwe 美特斯·邦威</div> <div>一家温州传统服装企业</div> <div>2016年亏损超一亿，近3年关闭1500余家门店。</div>	<div>ZARA</div> <div>一家欧洲时尚服装企业</div> <div>2016年上半财年Zara销售取得13%的增长。</div>	<div>韩都衣舍 HANDU.COM</div> <div>一家山东网络服装企业</div> <div>2016年双十一销售额达3.62亿，同比增长30%。</div>

## 三家企业的创新模式比较

✓传统模式，新品从设计到销售一般3-6个月	✓线上线下结合，新品从设计到销售一般15-30天	✓利用网络整合资源，新品从设计到销售7-15天
✓线下店铺分销，但无法快速掌握门店销售及用户反馈情况	✓线下店铺分销，但自建信息化手段，能够掌握当天销售及用户反馈情况	✓通过大数据，实时掌握各品类销售及用户反馈情况
✓供应链只能适应大批量、的生产，库存压力大	✓供应链能够适应中等批量，库存压力小	✓供应链能够适应小批量生产，具有7天生产、3天补货的能力，零库存

# 大数据推动政府治理变革（1）

**大数据正在成为政府推动社会民生服务普惠化、精准化、便捷化的重要手段。**

**由于信息缺乏和不对称所导致的民生服务小众、粗放、繁冗的缺陷在一定程度上得到改善。**

例如，大数据将使得全面衡量个人信用水平变得可能，从而可以使得金融服务面向更广泛的人群。经过大数据挖掘，精准营销、精准医疗、智能交通等民生服务的新模式正在成为现实。共享的数据将逐步打通各垂直行业间信息孤岛的现状，改善不同服务需要对相同数据进行重复操作的繁冗。利用数据创新性地实现资源优化配置，将推动社会服务的效率大大提升。





# 大数据推动政府治理变革（2）

## 大数据正在成为完善国家治理体系和提升治理能力的关键路径。

大数据推动了部门间、政府与民众之间的信息互联，加速消除信息孤岛现象。运用大数据技术开展智慧治理、数据决策、风险预警、智慧城市、智慧公安、舆情监测等，可以助力实现宏观调控科学化、政府治理精准化、商事服务便捷化。利用数据创新政府决策方案，可以大大提升公共服务能力。



例如，联合国“全球脉动计划”就是用大数据对全球范围内的社交网络数据和文本信息进行实时分析监测和“情绪分析”，可以对疾病、动乱、冲突提供早期预警。

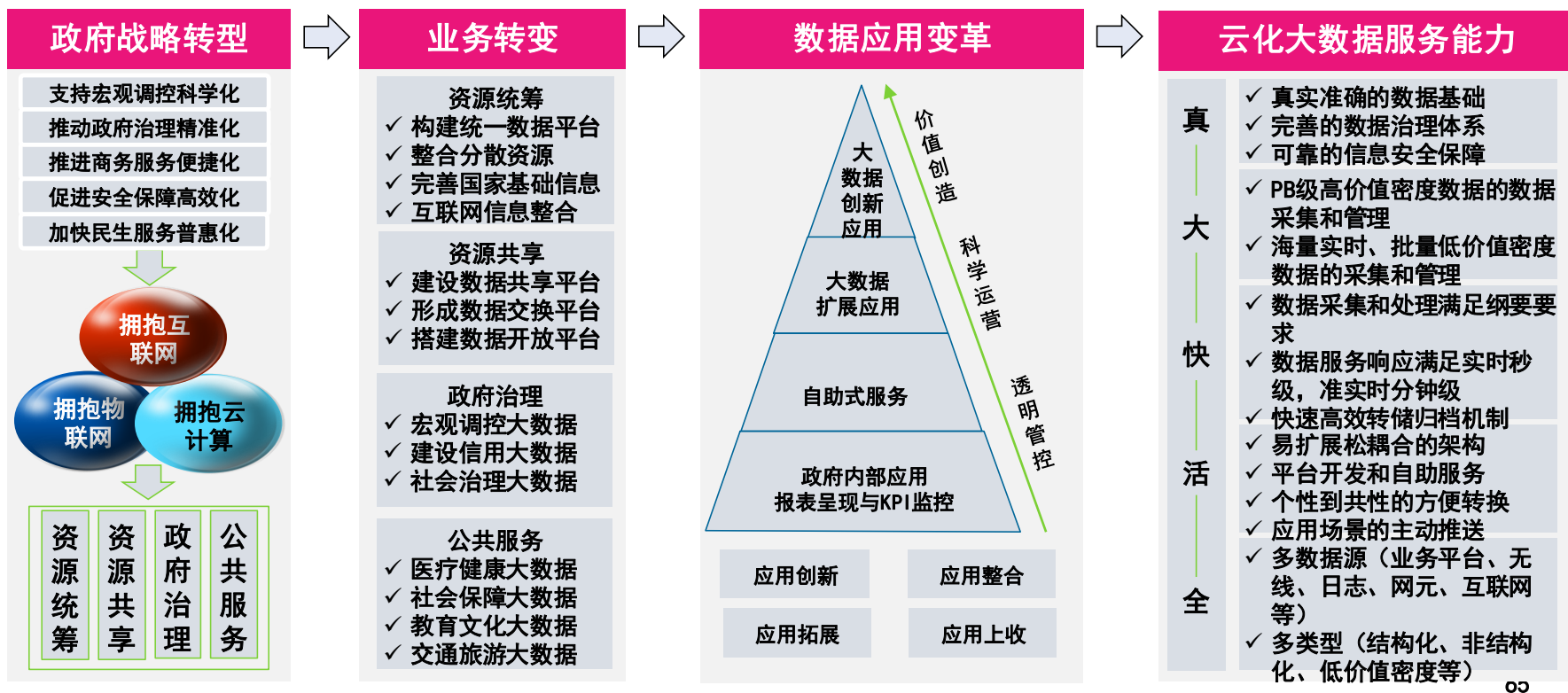
---

- **政府大数据建设的推进步骤与目标**



# 政务大数据的建设驱动力

- 国家数据战略要求实现政府业务转变和数据应用变革，进而驱动政务大数据建设和云化大数据服务能力提供
- 政务大数据实现政府内外部全面、海量数据的加工，以灵活、高效的方式提供真实准确的数据服务



# 当前政务大数据建设模式与问题

## 物理设备集中，应用分散建设

实现：IT 物理资源集中和管理。

问题：应用分散未集中化，基本上无法发挥“大数据”的作用。

## 应用集中建设，但未真正融合打通

实现：IT 物理资源集中和管理，基础信息化系统集中建设。

问题：应用未在统一平台上打通，大量功能重复建设，基础数据多点建设与数据不一致，无统一标准，扩展性差，应用平台很难实现高效协同，“大数据”作用不明显。

## 打通数据，统一标准，融合业务（推荐）

- ✓ 实现基础资源+数据资源+可共享业务能力集中化建设。
- ✓ 真正实现复用、节约投入成本。
- ✓ 做实平台层，共性能力由平台层统一提供，真正实现“厚平台，薄应用”。
- ✓ 统一标准和规范，全局用“一套语言”说话，打破信息孤岛，实现信息共享，解决数据不一致问题。

# 政务大数据建设步骤——①数据采集

## 数据采集问题

### 1、基层数据采集困难

- 需求太多：上级需求单位多、采集压力大
- 标准太乱：采集口径不同、采集标准不统一
- 管理太少：没有管理规范，没有数据留存要求，存在数据造假

推动基层跨部门  
共享校核



**“顶上千根线，下面一根针”**  
例如一个居委会，往往需要应对公安、民政、统计、劳保等各个部门的采集需求。在管理不规范的情况下，很难保证数据采集的质量。

### 2、属地化数据管理困难

信息系统纵强横弱。区域机构难以留存数据，难以建设区域统一的数据平台。基层数据难以形成体系。

例如：天津滨海新区爆炸案发生后，滨海新区不能第一时间调取爆炸区相关数据

### 3、历史珍贵数据不够重视

### 4、实时数据获取困难

# 政府大数据开放与共享的概念

## 开放

指数据供不特定第三方免费使用、再利用、再分发

## 共享

指数据在一定的条件下与指定的第三方进行共同使用

由于政府数据的**高价值和公益性**，政府数据的开放和共享格外受到关注。政府数据开放共享也成为数据开放共享的“排头兵”。



**政府数据开放**，是指行政机关面向公民、法人和其他组织提供政府数据的行为。

**政府数据共享**，是指行政机关因履行职责需要使用其他行政机关的政府数据或者为其他行政机关提供政府数据的行为。

# 政务大数据建设步骤——②数据开放

## 政策制约

- 政府数据开放缺乏统一政策文件，**无规可循**。
- 政府开放数据，助力社会创新的**机制缺乏**，动力不足。

## 基础制约

- 政府数据开放的**信息资源基础弱**，数字化不足，水平不一。
- 缺乏**集中统一**的政府数据开放的技术渠道。。



## 健全内部机制

- **发布机制**——初期采用有限的“白名单”，逐步向“黑名单”发展
- **公示机制**——第三方评估发布结果及排名
- **KPI机制**——增加在电子政务考核中占比
- **数据资产化管理机制**——全生命周期

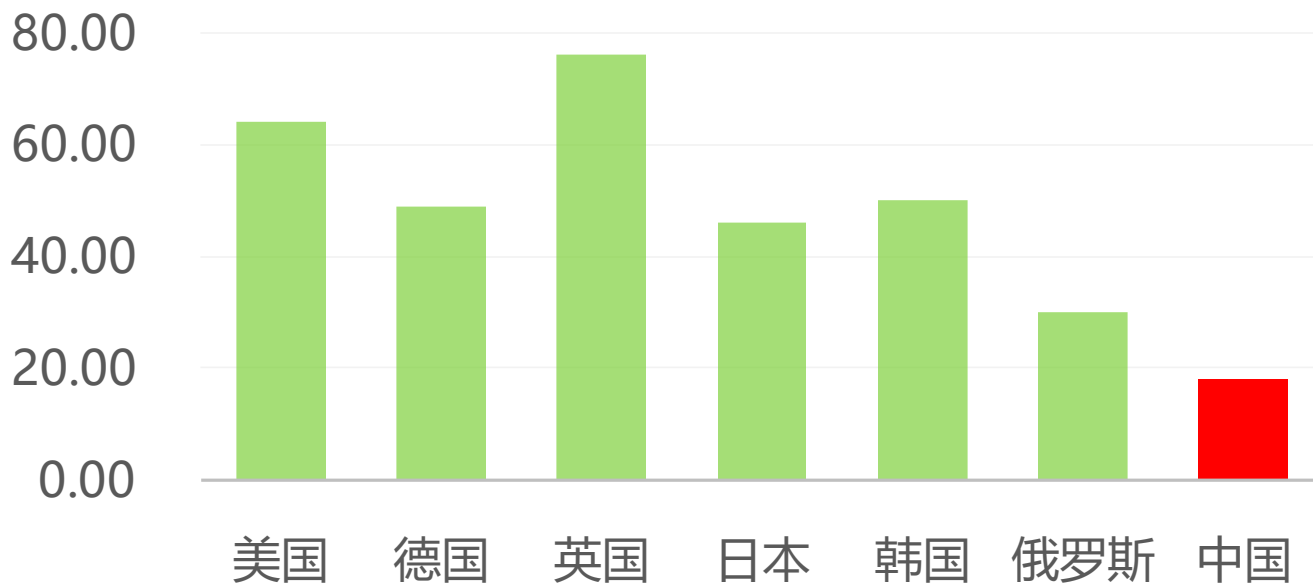
## 发挥外在动力

- **需求反馈机制**——企业需求+民众需求，支持政府与企业“资源置换”
- **社会资本参与机制**——政府采购服务、/数据、政企交换数据、社会众包等，
- **支持数据创新**——支持数据大赛，跨部门、领域数据融合项目作为重点

# 我国数据资源开放水平仍然不高

- 标准规范不健全，
- 尚未建立起兼顾安全与发展的数据开放
- 管理和信息安全保障体系。

OpenIndex2016年发布的评分显示，我国信息资源综合开放评分为18分，与主要发达国家的差距明显。



**数据来源：**  
OpenIndex网站于2016年公布的信息资源综合开放报告，根据其收集的数据对政府信息公开共享程度进行评分。

# 政府数据开放主要问题

## 不愿开

**信息资源是独家垄断资源**——开放后担心部门权利削弱、经济利益受损，并可能暴露部门业务问题。

## 不敢开

**尚缺乏保障数据开放的配套制度**——缺少具有可操作性的强制性规定，与政府信息公开、保密法、档案法等相关法律法规衔接不到位，各地对政府数据开放的范围和潜在风险存在“后顾之忧”。

## 不能开

**数据基础不牢**——公共部门尚未建立一套完整的数据资源采集、管理、加工和开发利用的体系，很多信息资源缺乏数字化，数据资源多头采集、重复建设、成本高昂，很多数据无人维护、不具有可持续性，数据的质量和准确性也存在问题，有哪些数据资源也不清楚。

## 不会开

**开放质量不高，可利用性差**——网站的建设和维护问题，网站数据的质量问题（可机读性差，数据更新频率不高，数据互动性差等）

**网站缺乏标准化**——增加使用者成本。



# 全国政府数据开放平台数量迅速增长

全国20多个数据开放平台开通运行，涉及14个省份。

省级	北京市	北京市政务数据资源网	2012年	共38家单位17个领域的296项数据集
	上海市	上海市政府数据服务网	2011年	共41个部门涉及11个领域的803个数据集
	浙江省	浙江政务服务网-数据开放频道	2014年	浙江网上政府24小时在线，650项数据
	贵州省	贵州省政府数据开放平台	2014年	共110个数据资源
	重庆市	重庆大数据平台-开放数据频道	2015年	共74个数据集
副省级市	武汉市	武汉市政府公开数据服务网	2015年	共48个政府部门755项数据
	深圳市	深圳政府在线数据开放平台	2016年	包括福田区、罗湖区等政府数据



# 政务大数据建设步骤——③数据共享

## 政策制约

- 政府数据共享缺乏全国统筹机制，**协同不足**。
- 政府共享数据，保护部门利益的**本位思想较重**。

## 基础制约

- 数据共享的交换机制导致**数据不落地**，难以满足大数据对多维度数据的需求。
- 缺乏**中心交换**的政府数据共享的技术渠道，没有数据融合的统一平台。

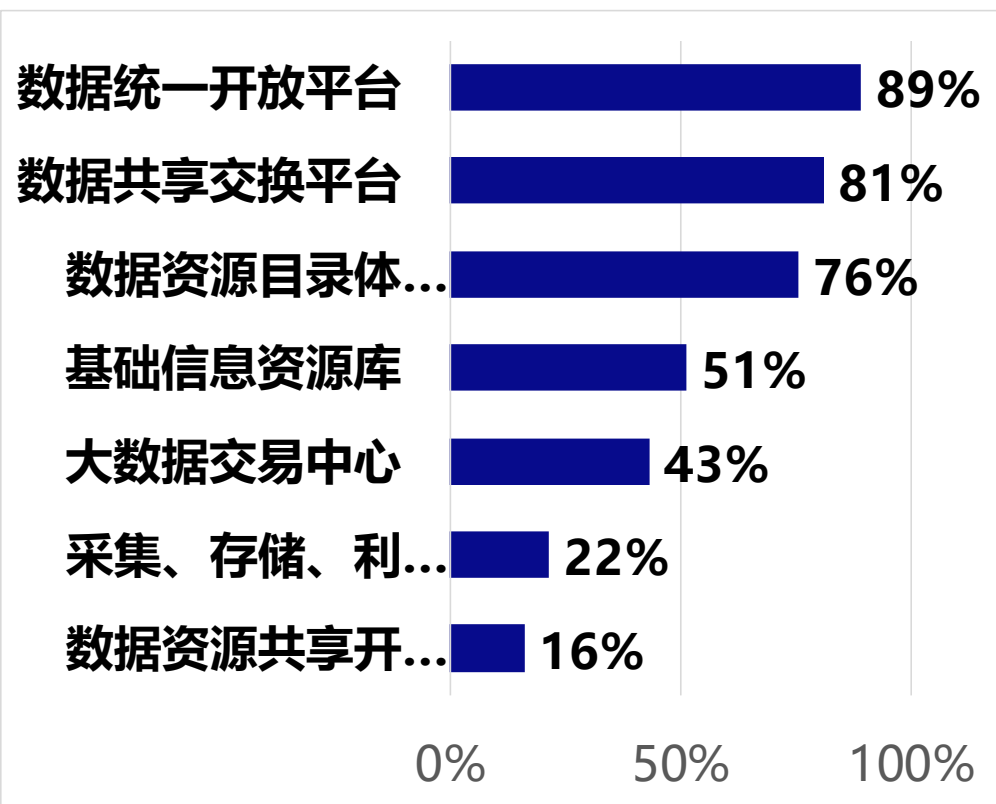
**核心：**完全共享数据应由统一平台来管理，尽量扩大全面共享数据范围



- 1、**数据分级/分类机制**：制定数据分级分类管理机制，对不涉及秘密和敏感内容的信息要在政府内部全面共享；不宜全面共享的数据应根据各部门履职需要及数据不同级别，按需定向共享。
- 2、**负责人制度**：指定部门数据共享专职负责人，审核可以共享的数据，制定数据需求和共享计划

# 数据开放共享成为各地方发展大数据的首要工作

各地纷纷加速数据共享、开放和流通，激活大数据产业新动能



- ❑ 截止到2017年1月底，16个省、21个市发布了大数据发展规划
- ❑ 80%以上的省市提出开展数据共享开放平台建设，强调政府数据的统筹协调管理
- ❑ 进一步完善统一的人口、法人、和空间地理等基础信息资源库
- ❑ 制定优化《政府数据资源共享与开放目录》和大数据开发利用标准

# 各地推出机制，保障数据开放共享稳步推进

- 建立**目录清单制度**成为各地推动公共数据开放的实现基础和政策抓手。



## 建立目录清单制度

- 目前主要采用**正面清单**形式
  - **上海**：自2014年起要求各部门每年围绕重点领域制定并公布开放清单
  - **贵州**：出台《贵州省大数据发展应用促进条例》，数据开放引入负面清单
  - **广东（珠三角）**：开展目录清单的梳理和建设工作



## 加大资金支持力度

- 目前专项资金支持模式尚未明确
  - **上海**：信息化发展专项资金支持方式
  - **北京（京津冀）**：提出要加大财政资金对数据网站的支持力度
  - **河南**：提出了公共信息资源开放资金计划



## 探索考核机制

- **上海**：纳入政府信息公开考核，确保数据开放工作常态化，建立信息化建设与政府信息资源开放工作联动机制，进行绩效考核



## 健全安全保障制度

- **深圳**：加快数据安全保护立法
- **北京（京津冀）**：提出建立公共信息资源开放利用分级认证制度，有效规范应用管理

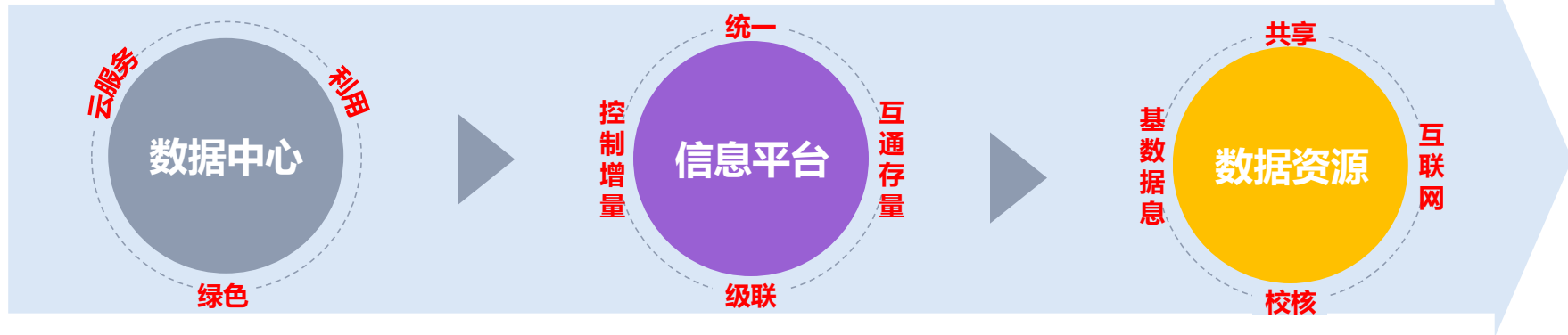


## 加强组织协调

- **上海**：建立政府信息资源开放工作领导协调机制，指定推进部门，区县、部门共同参与
- **广东（珠三角）**：成立省、市大数据管理局，统筹数据资源管理
- **贵州**：成立数据资源管理机构

# 政务大数据建设步骤——④统筹汇聚

引导数据中心、信息平台 and 各类数据库统筹集约化建设，实现全市级大数据平台



- 向云服务迁移，采用“租建结合方式”
- 绿色集约型数据中心
- 建立数据中心资源审计监管平台

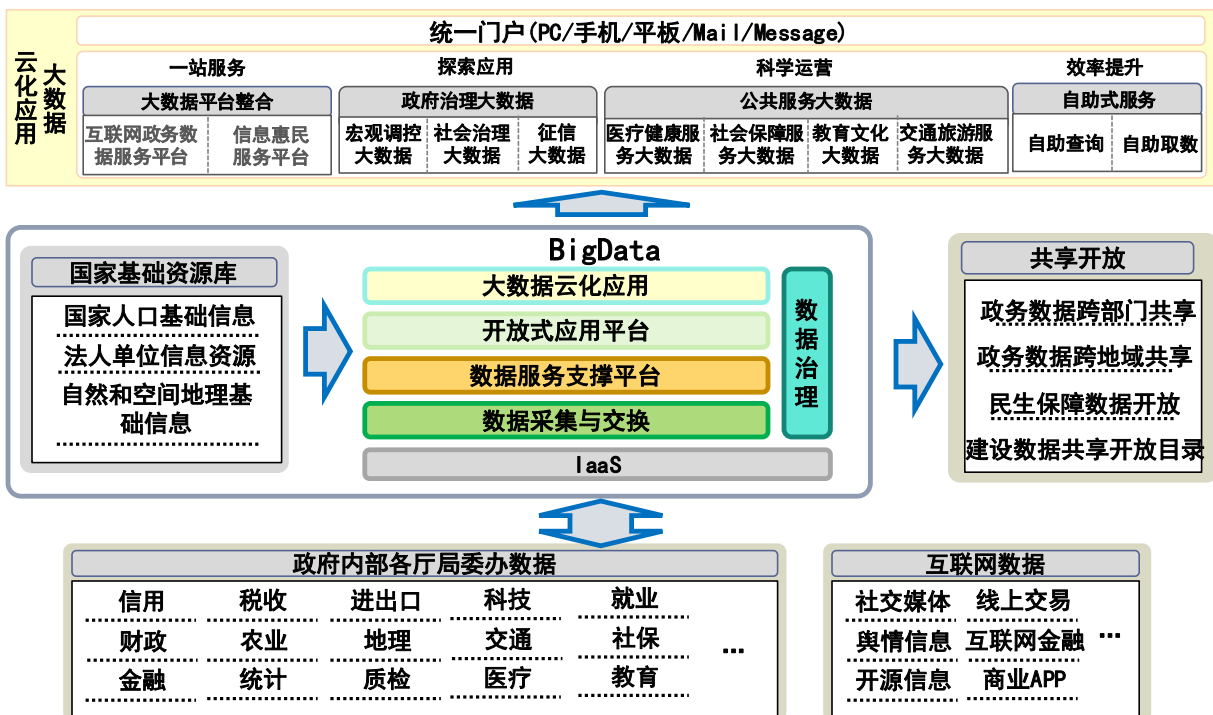
- 统一的信息服务平台
- 联通现有各级平台，形成“多级连接、逻辑集中”的服务平台
- 严格控制增量审批

- 建设基础数据资源
- 融合社会数据资源
- 采集互联网数据资源
- 数据应反垄断法

- 1、由于数据关联分析的价值难以预估，**跨部门的数据融合**缺乏动力
- 2、**增量评审**是推动抓手，新建信息化系统将以数据汇聚作为评审条件

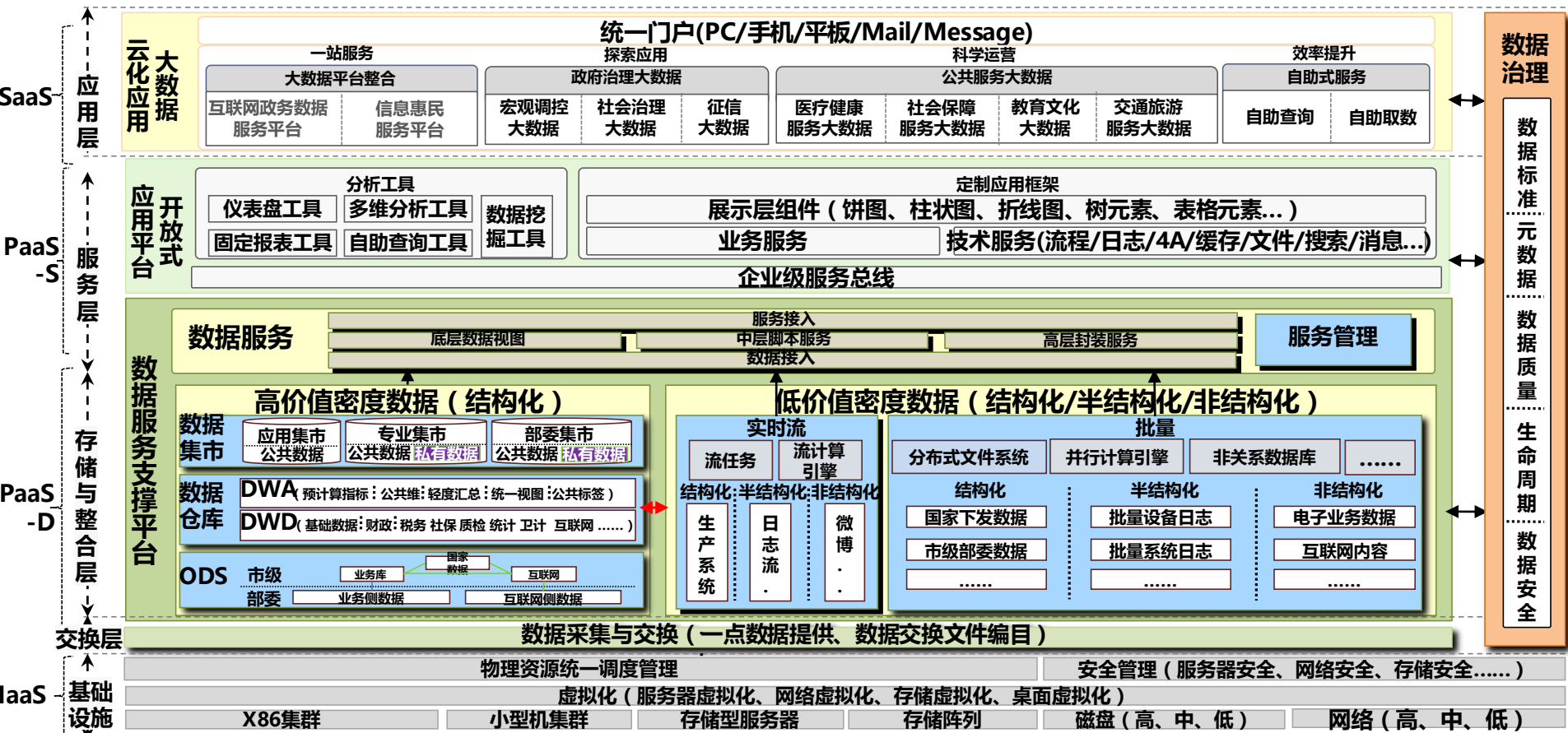
# 政务大数据云平台建设目标

- 建立能够支撑政府大数据战略转型的**大数据云化应用体系**，以及数据资源整合后，满足各厅局委办海量用户数据服务需求和多厂商合作的能够个性化定制、灵活组装的**应用开发平台**
- 建立支撑大数据应用体系，满足数据大集中要求的**数据服务支撑、数据采集交换和数据治理三大平台**
- 建立一套支撑政府大数据应用目标的**组织管理机制**，在数据集中时，保证服务下沉为政府各部门提供大数据支撑能力









- 实现六大组件：大数据云化应用、开放式应用平台、数据服务支撑平台、数据采集与交换、IaaS层、数据治理。
- 提供六个一点能力：一点采集、一点数据交换、一点加工转换、一点数据存储、一点分析挖掘、一点提供服务。
- 兼顾集中管控与个性服务
  - 数据共享，集中管控，有效应用，产业发展
  - 一站服务，科学运营：政府内部应用通过一站服务提供，探索政府治理类大数据应用，科学运营公共服务大数据，并通过自助式服务提升办政效率。
  - 实现系统上收和服务下沉：集中整合各部委数据，数据服务能力通过数据开放、数据应用、自助服务下沉至一线

# 政务大数据总体方案



# 大数据技术平台的选型需要公平的衡量标准

## 开源软件→ 商用产品

	可用性：包括系统的稳定性和高容错性能力等
	架构兼容性：与原有系统的兼容性（SQL接口），与开源架构的兼容性
	性能：高吞吐量和性价比，水平线性扩展能力
	安全性：隐私数据的加密，严格的权限管理，完善的数据保护
	易用性：快速的集群部署以及扩展，自动调优，自动化平台升级
	灾备能力：业务连续性和监管合规



# 大数据基础平台产品逐步走向成熟

运维管理	可用性	功能	兼容性	安全	多租户	扩展性
自动化部署	Namenode主节点失效恢复	数据导入	ODBC兼容性	认证	租户管理	集群动态扩展
资源监控	Namenode备节点失效恢复	SQL任务能力	JDBC兼容性	授权	资源管理	集群动态收缩
作业监控	Datanode节点失效恢复	NoSQL数据库	SQL支持度	加密	资源隔离	
集群操作	HMaster节点失效恢复	机器学习	传统数据库同步	审计	权限管理	
故障管理	RegionServer节点失效恢复	流处理能力	跨不同数据库表关联操作			
日志管理	HDFS备份恢复					
配置管理	HBase备份恢复					
权限管理	运维管理节点失效及恢复					
用户管理						
无宕机升级						

大数据产品基础能力认证包括七大项：功能、运维、多租户、可用性、安全、兼容性、扩展性，总共**38**项测试用例

# 大数据基础平台能力评测——量化技术实力

数据中心联盟 《大数据基准测试2.0》 **先后四轮测试**

全球首次在**统一的测试平台**上进行测试

国内**第一个工业级大数据**基准测试活动，20多家国内顶级大数据商业发行版厂商

国内**第一个多个任务负载**的大数据基准测试活动，覆盖性能、容错性、安全与运维

国内外**第一次对大数据公有云服务**进行测试

中国电子科技集团公司第二十八研究所

中兴通讯技术有限公司

成都四方伟业软件股份有限公司

杭州泰一指尚科技有限公司

北京中联润通信息技术有限公司

航天信息股份有限公司

联想（北京）有限公司

深圳前海信息技术有限公司

中国华戎科技集团有限公司

28th鲸云大数据存储分析平台软件[鲸云大数据平台]V2.0

大数据平台DAP V2.0

SDC大数据基础平台V3.0

泰一指尚DMP数据聚合与管理服务平台V1.0

易云汇智大数据平台V1.0

航天信息大数据分析平台V1.0

联想企业级数据分析平台[LEAP系统]V3.2

前信闪电大数据系统 V1.0

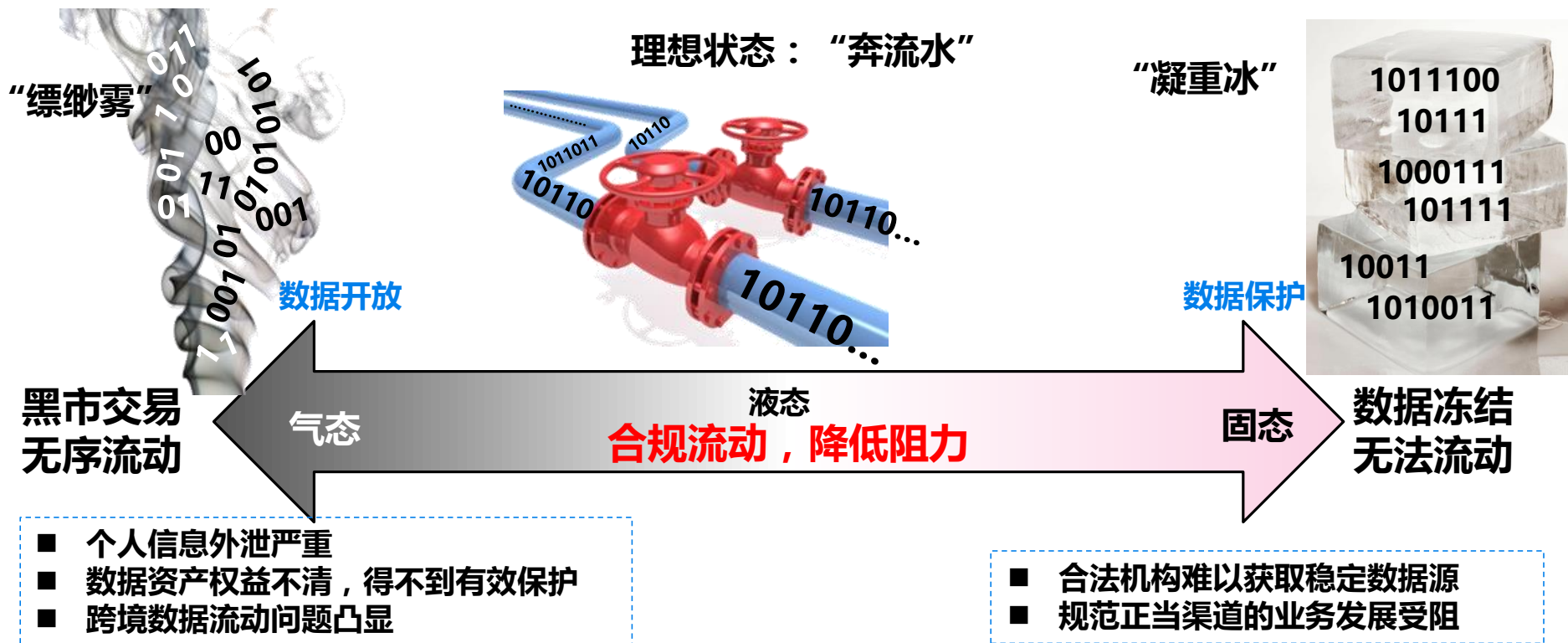
华擎大数据平台 V2.0

# 大数据基础平台的技术现状

企业	无宕机升级	运维管理节点HA	SQL支持度	跨数据库关联	多租户资源隔离	多租户权限
中电28所	大版本升级	√	93条语句	√	√	√
中兴	小版本升级	√	99个语句	√	√	√
中联润通				√		
泰一指尚	大版本升级	√	99个语句	√	√	√
航天信息			10个语句		√	√
联想	小版本升级	√	99个语句	√	√	√
前海信息	小版本升级	√	10个语句	√	√	√
四方伟业	小版本升级		10个语句	√	√	√
华戎		√	99个语句	√	√	√

# 政务大数据建设步骤——⑤流通交易

## 一、数据流通交易面临两个极端



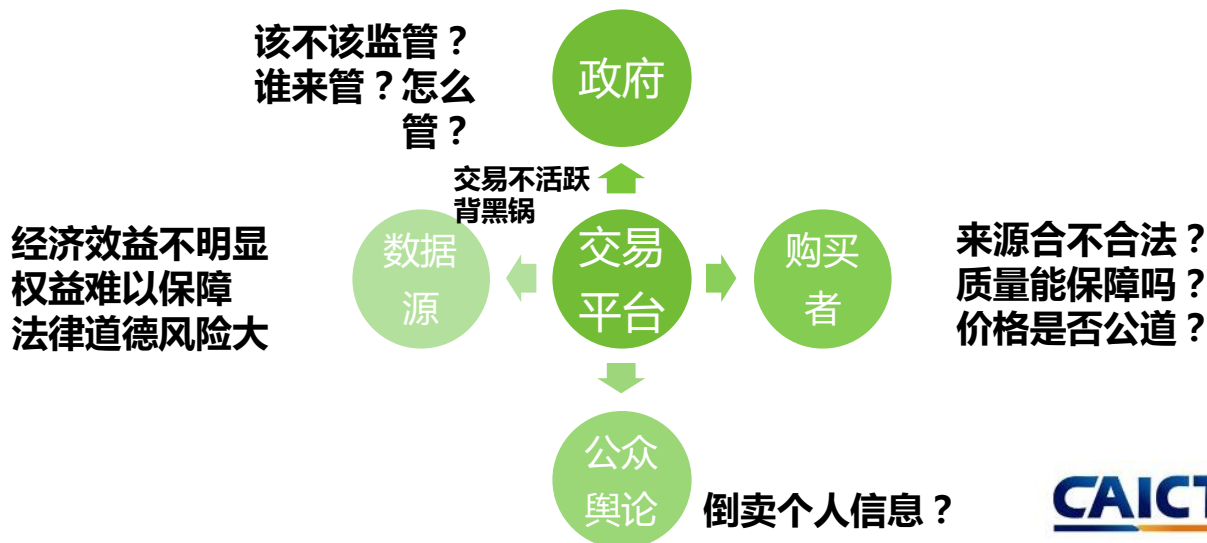
# 数据流通高速发展，各方利益相互交织

## 地方大数据交易平台建设进入井喷期

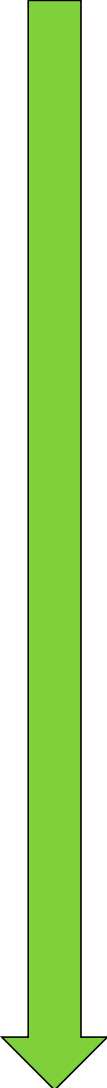


截至2017年5月，全国已建成18个以上大数据交易场所，到2020年，全国数据交易平台数量预计达到30个

北京	上海	浙江	贵州
重庆	深圳	广州	武汉
贵阳	无锡	湛江	.....



# 解决思路1：采用完善数据流通安全措施



**技术安全措施：**用于生成特有身份识别信息的算法信息在数据去身份化之后及时销毁，使去身份化不可被撤销。

**业务流程控制：**单位应当制定去身份化流程以及去身份化技术实施方案，并定期审查和更新，以确保流程和方案适当且符合当前的技术水平和行业实践。

**数据分级制度：**建立严密的内控流程，根据数据的敏感层度进行数据敏感分级，针对不同分级，在使用数据前，需要经过相应敏感级别对应所需的审批流程。

**完善管理制度：**制定对于身份关联信息或去身份化信息的管理制度，规范员工行为，并设置专门人员定期对于业务流程进行审核，尽量将数据的使用控制在具有安全防范措施的平台内部，避免敏感数据被调出后存在可能泄露的风险。

# 解决思路2：推进大数据法律制度建设

- 在大数据和“互联网+”背景下，结合我国数据产业发展和行业监管实践，大数据管理立法应该明确“安全与发展并重、促进与规范并举”的战略方向和遵循“积极利用、科学发展、依法管理、确保安全”的方针。

- 释放数据经济活力
- 数据商业利用

立法原则

- 保护数据主权与安全
- 保护个人隐私

## 制定整体性、基础性制度

- 确保数据安全性与个人隐私保护，强化数据收集、使用规则；
- 加强跨境数据流动管理，建立风险评估和安全审查机制；
- 明晰数据产权，抓紧确立数据交易规则；
- 推动出台政府数据开放法规政策

## 就重点问题制定针对性制度

- 高价值、高风险的重点数据：用户数据、用户个人信息。
- 涉及数据价值变动和法律关系变化的重点行为：数据开放、数据共享、数据交易、数据承接等。
- 与权益保护相关的制度：泄露通知等。



# 解决思路3：构建行业自律

由中央网信办、国家发改委、工业和信息化部指导，中国信息通信研究院发起，数据流通相关核心企业及大型互联网企业参与制定的“数据流通行业自律公约”已发布。

## 《数据流通行业自律公约》

公约2.0版  
架构

总则

数据  
权益

数据  
流通

数据  
应用

附则

公约1.0版  
内容

公约宗旨

1.遵守法律

2.合规采集

4.数据保护

3.流向管控

5.权益协调

6.服务提升

7.社会监督

新增内容

数据流通  
的定义

企业对数据  
的权益

强化流通环  
节风险评估

企业增强  
社会责任

公约的持续  
完善更新

完善内容

个人、企业  
的利益协调

第三方平台  
责任义务

推动技术标准  
与认证、审计

成立日常监  
督管理机构

# 解决思路4：标准化为数据流通提供保障

目前，数据中心联盟正在制订一系列数据流通相关标准，并委托中国信息通信研究院进行标准符合性测试。

安全管理 机制	数据使用 规范	身份识别	信息验证	数据转移 方式	授权方式 及流程	接口	审计
------------	------------	------	------	------------	-------------	----	----

征信类产品及服务	第一批测试已完成
----------	----------

金融风控类产品及服务	第一批测试已完成
------------	----------

数据流通中心数据处理及业务管理	第一批测试已完成
-----------------	----------

精准营销类产品及服务	标准制订中
------------	-------

位置信息类产品及服务	标准待制订
------------	-------

# 数据流通平台合规性现状

## 首批通过**征信服务类数据流通标准**符合性测试的企业



中国电信股份有限公司云计算分公司



中国联通信息化事业部数据中心

## 首批通过**金融风控类数据流通标准**符合性测试的企业



四川享宇金信金融服务外包公司



北京通付盾数据科技有限公司

## 首批通过**数据流通中心数据处理及业务管理标准**测试的企业



上海数据交易中心



京东万象



安徽大数据交易中心

---

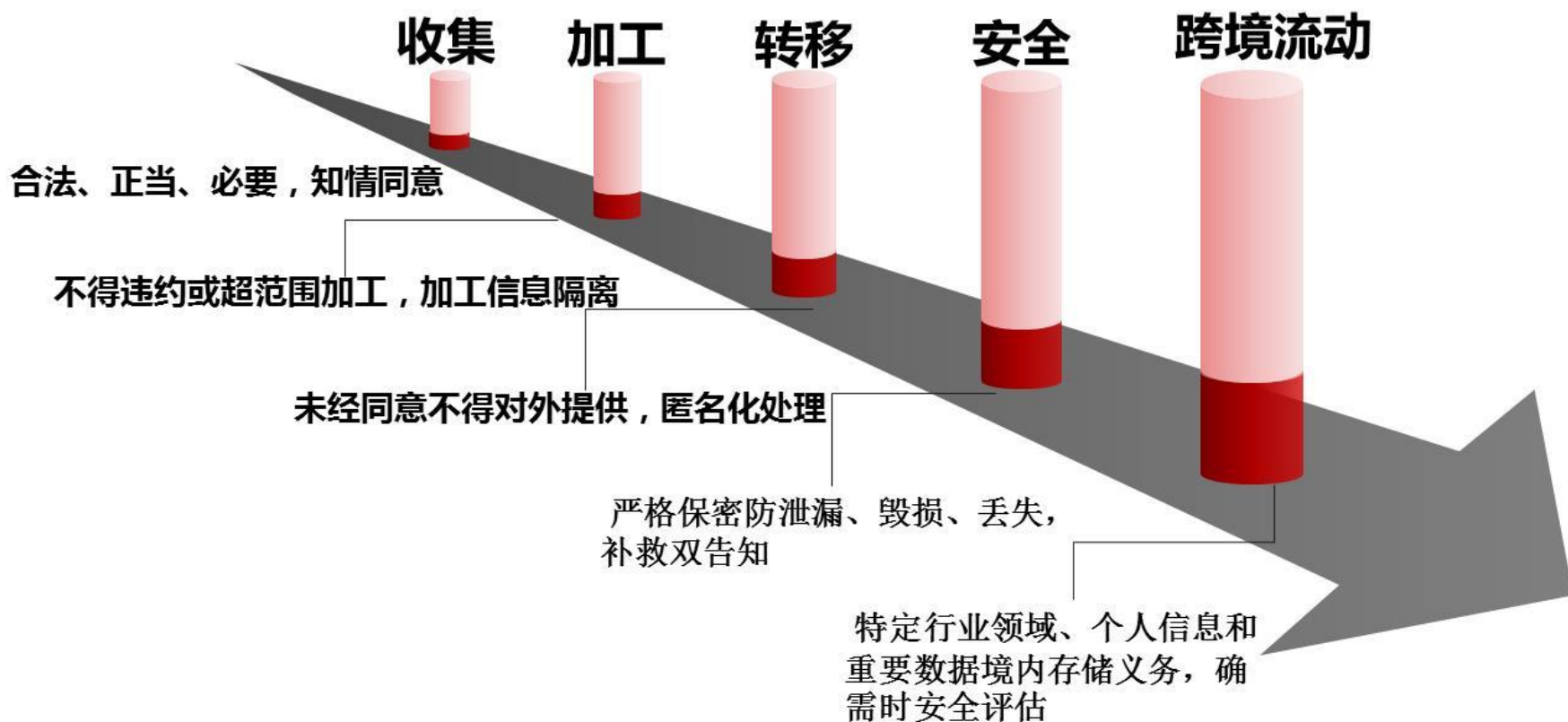
## 四、大数据合规与《网络安全法》

- 推进大数据发展的合规性
- 《网络安全法》的出台及应用

---

- **推进大数据发展的合规性**

# 大数据存在合规性要求



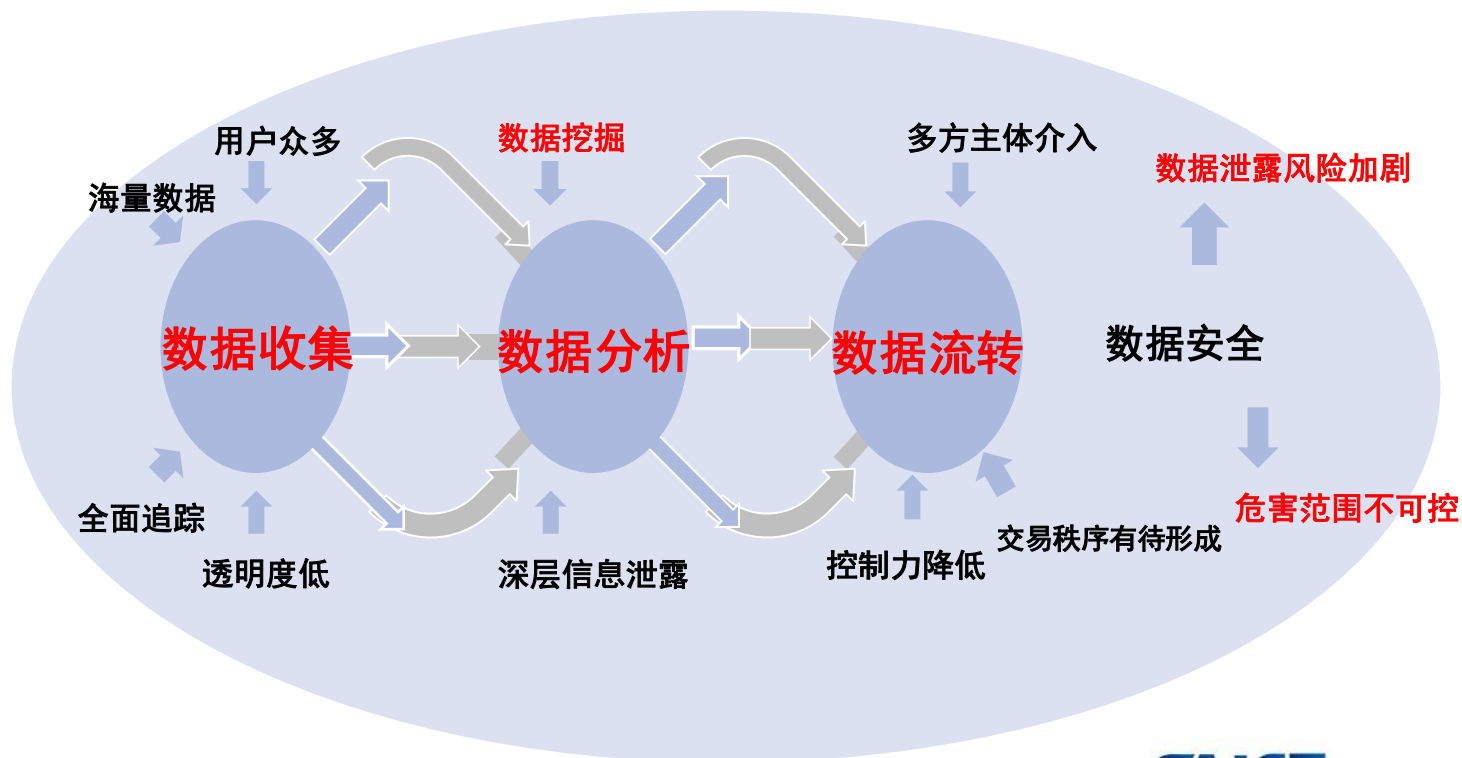
# 我国关于大数据安全保护方面的有关文件

- **《促进大数据发展行动纲要》提出：**
  - 在依法加强安全保障和隐私保护的前提下，稳步推动公共数据资源开放。
  - 明确数据采集、传输、存储、使用、开放等各环节保障网络安全的范围边界、责任主体和具体要求，切实加强对涉及国家利益、公共安全、商业秘密、个人隐私、军工科研生产等信息的保护。
- **《大数据产业发展规划（2016-2020年）》提出：**
  - “坚持发展与安全并重，增强信息安全技术保障能力，建立健全安全防护体系，保障信息安全和个人隐私。加强行业自律，完善行业监管，促进数据资源有序流动与规范利用”
- **《中华人民共和国网络安全法》提出：**
  - “国家鼓励开发网络数据安全保护和利用技术，促进公共数据资源开放，推动技术创新和经济社会发展”



# 大数据环境下个人数据保护形势严峻

大数据技术的广泛应用为个人数据保护带来新的挑战，集中体现在**数据收集**、**数据分析**、**数据流转**等环节，及增大**数据安全**及隐私侵害**风险**等方面。



# 数据跨境流动带来安全风险

## 数据泄露风险

境外当地数据保护水平不足，数据泄露

境外当地执法部门执法强制披露数据

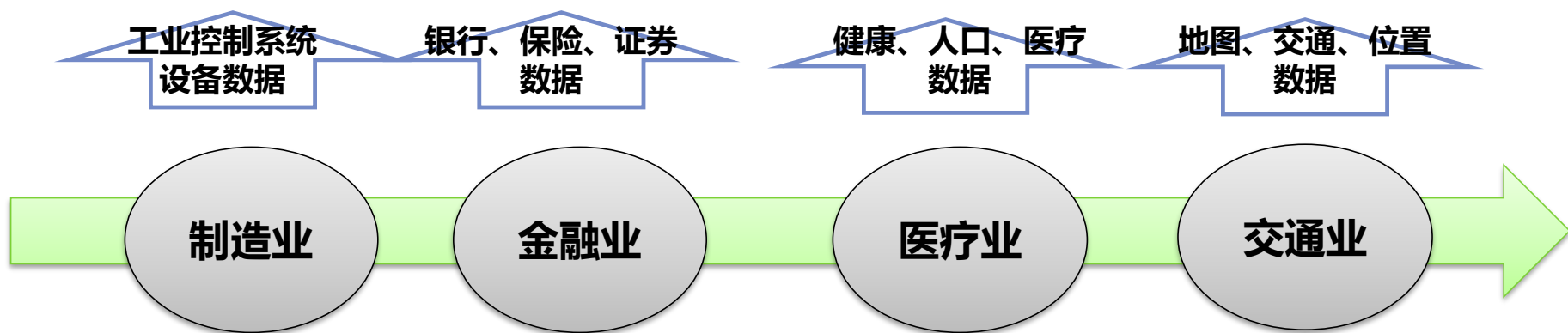
境内侵权行为维权困难

## 丧失数据资源优势

各国信息基础能力差异导致收集、处理、运用数据能力差异

美等强国掌握各国大量关乎经济社会及国家安全的**关键信息**

部分国家形成更强的**信息资源优势**和**战略控制能力**



---

- **《网络安全法》的出台及应用**

# 《网络安全法》立法进程

- 我国近年来加快了网络安全方面的立法步伐。2015年7月1日，十二届全国人大常委会第十五次会议通过《国家安全法》，明确要求“维护国家网络空间主权、安全和发展利益”，“实现网络和信息核心技术、关键基础设施和重要领域信息系统及数据的安全可控”，网络安全成为国家安全的重要组成部分。
- 《网络安全法》如期通过。在《国家安全法》的大框架之下，为解决我国网络安全领域存在的突出问题，以制度建设提高国家网络安全保障能力，掌握网络空间治理和规则制定方面的主动权，切实维护国家网络空间主权、安全和发展利益，顺应国内外日益复杂的网络安全态势，迫切需要在国家层面制定一部统一的网络安全基本法。在此背景下，2015年7月6日，全国人大网公布《网络安全法》（草案），向社会公开征求意见。2016年7月5日，全国人大第二次公布《网络安全法》（草案），向社会公开征求意见。
- 2016年11月7日全国人大常委会正式出台《网络安全法》，自2017年6月1日起施行。

# 《网络安全法》主要内容

序号	主要制度内容
(一)	网络空间主权和网络安全战略 (第1、4条)
(二)	网络安全监管体制 (第8条)
(三)	网络安全等级保护制度 (第21条)
(四)	网络关键设备和网络安全专用产品安全认证检测制度 (第23条)
(五)	网络实名制的网络身份管理制度和网络可信身份战略 (第24条)
(六)	合法侦听协助制度 (第28条)
(七)	关键信息基础设施运行安全保护制度 (第31-34条、第38条、第39条)
(八)	关键信息基础设施网络产品服务安全审查制度 (第35条)
(九)	关键信息基础设施重要数据境内留存制度 (第37条)
(十)	个人信息保护制度 (第40-44条)
(十一)	网络信息管理制度 (第46-49条)
(十二)	重大突发社会安全事件的网络通信临时限制制度 (第58条)

# 逐步完善网络安全立法

## 《网络安全法》是网络安全立法的新起点而非终点。

### ● 为建设网络强国提供制度保障：

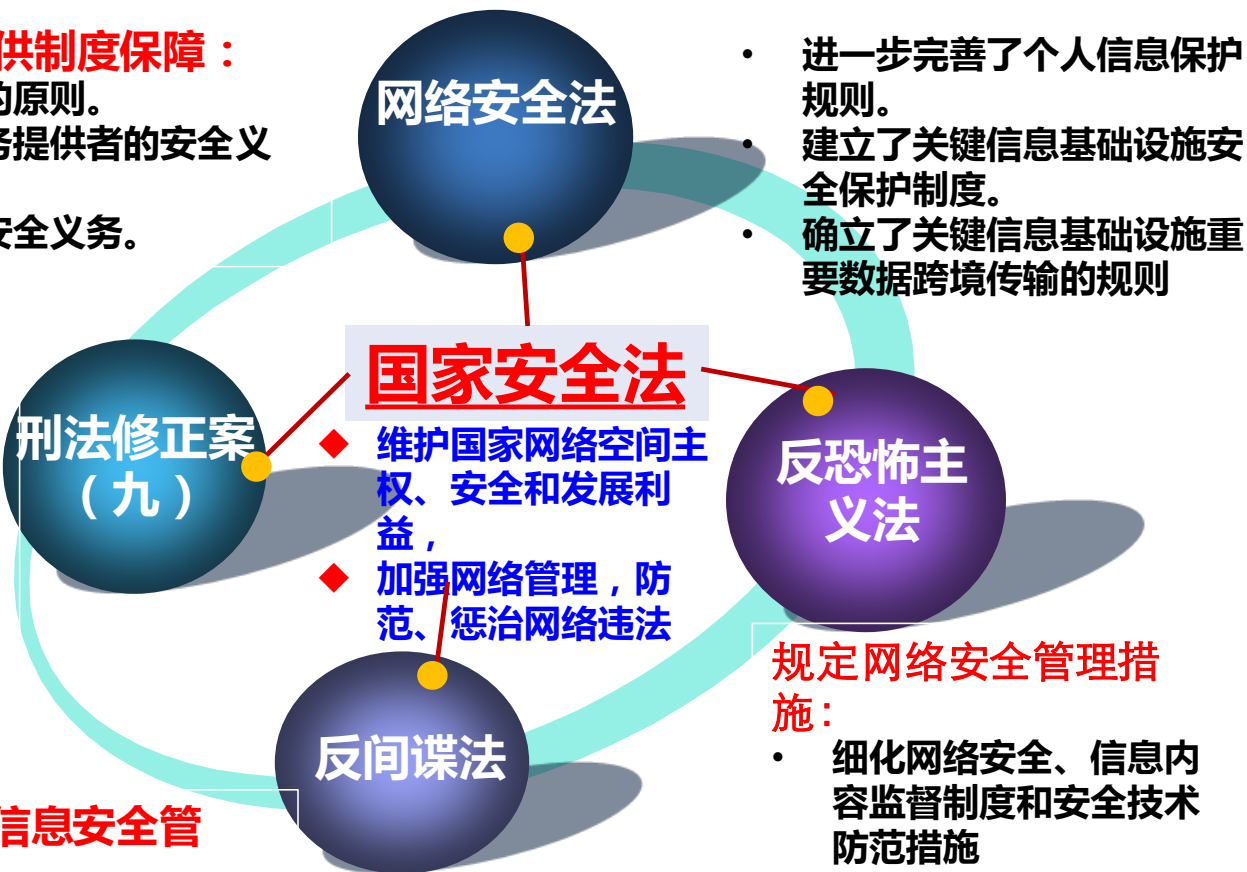
- 明确了网络空间主权的原则。
- 明确了网络产品和服务提供者的安全义务。
- 明确了网络运营者的安全义务。

### ● 护航信息网络安全：

- 加强对公民个人信息的保护
- 明确ISPS履行网络安全管理的义务
- 完善网络犯罪规定
- 增加编造、传播虚假信息犯罪的规定

### ● 明确间谍行为强化信息安全管理：

- 第十三条：国家安全机关因反间谍工作需要，可以依照规定查验有关组织和个人的电子通信工具、器材等设备、设施。



# 《网络安全法》确立了个人信息基本保护制度

规定了相关主体的个人信息保护义务，为大数据“划底线”

## ◆ 对于网络运营者等企业主体而言，要求：

- ✓ 在收集使用个人信息的时候遵守《网络安全法》规定的基本原则（第41条）；
- ✓ 未经被收集者同意，不得向他人提供个人信息（第42条）；
- ✓ 个人信息的更正和删除义务（第43条）
- ✓ 应当建立网络信息安全投诉、举报制度，及时受理并处理有关网络信息安全的投诉和举报（第49条第1款）；
- ✓ 并积极配合网信部门和有关部门依法实施的监督检查（第49条第2款）。

## ◆ 网络安全监督管理部门及其工作人员，要求：

- ✓ 必须对在履行职责中知悉的个人信息、隐私和商业秘密严格保密，不得泄露、出售或者非法向他人提供。此外，任何个人和组织不得窃取或者以其他非法方式获取个人信息，不得非法出售或者非法向他人提供个人信息（第50条）。



# 《网络安全法》为大数据流通留下发展空间

《网安法》严格保护个人数据，但也通过“匿名化”等规定为大数据的流通提供了可能

## ◆ 在数据流通方面，

第四十二条虽然要求“未经被收集者同意，不得向他人提供个人信息”，但本款也规定了例外情形，即“经过处理无法识别特定个人且不能复原的除外”。个人信息数据匿名化处理是维护数据安全的有效手段，也是兼顾隐私保护和数据利用的有效路径。

## ◆ 在数据交易方面，

《人大决定》明确规定“不得出售”公民个人信息，但《网络安全法》第四十四条规定“任何个人和组织不得非法出售或者非法向他人提供个人信息”，也就是说，在一定合法情形下是可以出售和提供公民个人信息的，这无疑给符合规定的个人信息交易开了绿灯，为我国大数据产业发展提供了空间。

# 两高解释带来企业数据合规的新挑战



在《网络安全法》2017年6月1日正式生效前，最高人民法院与最高人民检察院于2017年5月9日发布《最高人民法院、最高人民检察院关于办理侵犯公民个人信息刑事案件适用法律若干问题的解释》（简称“《两高解释》”）及相关典型案例并将与《网络安全法》同时生效。

《两高解释》第一次对“侵犯公民个人信息罪”的犯罪要件做出了较为明确的规定，对于明确刑事责任的范畴和判断具有重要意义。

- 明确“违反国家有关规定”范围
- 明确未经同意提供公民个人信息构成犯罪
- “以其他方式非法获取公民个人信息”的概念扩宽

《两高解释》和《网络安全法》的生效，将大大提高企业在数据保护、网络安全保护等方面的合规义务

# 大数据企业《网络安全法》合规性指引（1）

## 1. 对个人信息数据进行脱敏处理可实现“合规”

在数据流通方面，第四十二条要求“未经被收集者同意，不得向他人提供个人信息”，但“经过处理无法识别特定个人且不能复原的除外”。**网络安全法把“个人隐私”和“数据信息”加以了区别式定义**，即“经过处理无法识别特定个人且不能复原的”的才可流通。大数据企业应在此条款的指导下，合法开展大数据应用。

# 大数据企业《网络安全法》合规性指引（2）

## 2. 明确信息收集和使用的目的、方式、范围可实现“合规”

《网络安全法》第四十一条规定“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”。因此，大数据企业在进行数据收集和使用时，需做到明示收集、使用信息的目的、方式和范围，并经被收集者同意。

# 大数据企业《网络安全法》合规性指引（4）

## 3.关键信息境内存储可实现“合规”

第三十七条要求“关键信息基础设施的运营者在中华人民共和国境内运营中**收集和产生的个人信息和重要数据**应当在境内存储。因业务需要，确需向境外提供的，应当按照国家网信部门会同国务院有关部门制定的办法进行安全评估；法律、行政法规另有规定的，依照其规定。”因此，大数据企业应更加关注自身数据的存储安全。一般情况下应确保收集和产生的数据在境内存储。

# 大数据企业《网络安全法》合规性指引（3）

## 4.重要系统和数据库容灾备份可实现“合规”

第三十三、三十四条要求“关键信息基础设施的运营者，应当确保其具有支持业务稳定、持续运行的性能”，并通过设置专门安全管理机构和安全管理负责人；定期对从业人员进行网络安全教育、技术培训和技能考核；**对重要系统和数据库进行容灾备份**；制定网络安全事件应急预案等方式履行安全保护义务。因此，对于关键信息基础设施的运营者来说，对重要系统和数据库进行容灾备份是必要且重要的。

---

## 结束语：促进大数据发展的建议



# (一) 加快数据开放共享，增强发展新动能

**我国数据开放刚起步**：仅北京、上海、无锡、青岛等地区分散开放了数据资源；  
2014年开放知识基金会数据开放评价中，我国第57（共97国）位

2014年 权威机构英国开放知识基金会 十大关键领域数据开放指数评价

Rank	Place	Transport Timetables	Government Budget	Government Spending	Election Results	Company Register	National Map	National Statistics	Legislation	Postcodes / Zipcodes	Pollutant Emissions	Score
1	United Kingdom											97%
2	Denmark											83%
3	France											80%
4	Finland											73%
5	Australia											72%
57	China											37%

## 问题

在交通、地图等十大关键领域，我国数据开放的整体水平低

## 举措

## 落实《大数据行动纲要》

优先开放高价值数据

——无争议、惠民生、易开放的数据先行开放

推动政府数据跨部门共享

——加快互联互通和信息共享，丰富信息服务

## （二）推动应用创新，提升政府治理新能力

### 推动应用创新，全面提升政府治理能力

#### 宏观调控

提升大数据分析技术在**风险预警、执行监督、宏观调控决策支持**等领域的应用水平

#### 保护环境

加强**环境数据**的实时观测和采集，**精准锁定污染源**，预测环境质量

#### 社会治理

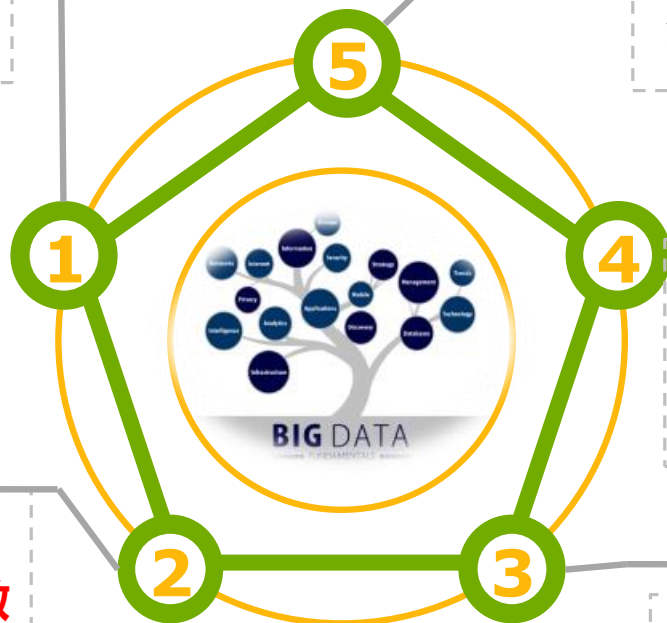
提高大数据分析技术在**治安管理、刑事侦查、公共危机管理、社区治理**中的应用水平

#### 公共服务

运用大数据技术开展**精准医疗、智能教育、智能交通、精准社保、精准扶贫**等

#### 市场监管

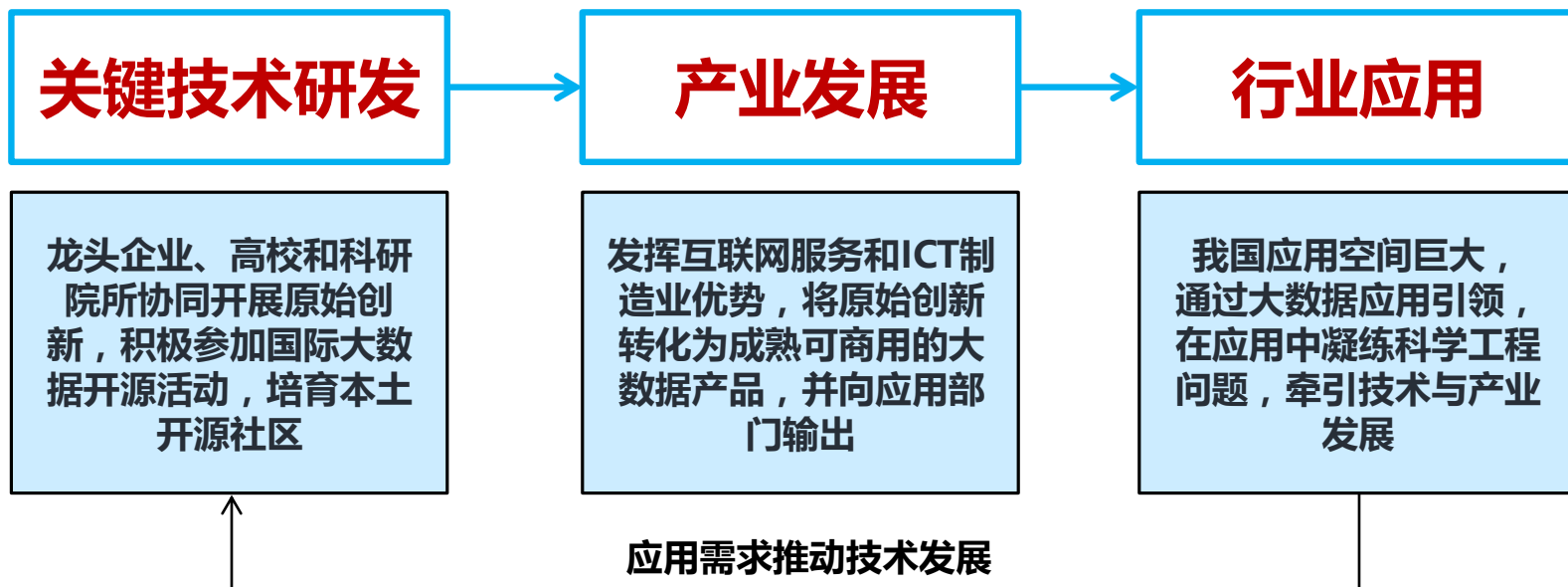
监测**市场运行**，汇集**市场数据**，分析**市场趋势**，推动市场监管智能决策和治理



## （三）加强关键技术研发，培育产业新生态

《纲要》：推进基础研究和核心技术攻关，形成大数据产品体系

- 着力突破核心技术**：海量数据存储、数据清洗、数据分析、可视化等
- 加强关键产品产业化**：存储软硬件、新型数据库、分布式计算平台、智能软件等
- 繁荣产业健康生态**：软硬件互相促进，资源-产品-服务良性发展的产业生态
- 以应用带动产业发展**：推动政府部门、重点行业及新兴行业应用



## （四）推动大数据创业创新，开辟增长新空间

数据成为创业创新的关键要素。通过发展一批大数据开放平台、大数据众包平台、大数据众创平台和大数据技术平台，全面助力创业创新。

### 培育大数据众包平台

鼓励大数据企业在数据的采集、计算、分析等环节积极应用众包模式，降低成本、激发大众创新

### 组建大数据技术平台

综合科研院所、行业组织、企业等多方优势，开展基础理论、关键技术、应用和工具的重点攻关

### 发展大数据开放平台

鼓励大平台、大企业向第三方开发者开放数据资源，鼓励数据应用的跨界创业创新

### 建设大数据众创平台

建设一批大数据众创平台，专业孵化助力创业创新

给你所需 给你所想  
**百度数据开放平台**

百度数据开放平台是百度基于海量数据资源，面向开发者、企业、机构，提供合法数据资源开放、数据清洗和数据应用开发、数据应用开发成果“数据应用”的开放平台。



## （五）加强数据安全 管理，营造发展新环境

### 规范数据的收集与使用行为

- 美欧等发达国家在隐私与数据保护立法基础上，针对具体新技术业务，进一步发布数据收集与利用指南，指导行业合规收集与使用数据；
- 我国应借鉴国际制度经验，加强对数据利用的合规规范和指导

### 推动建立跨境数据流动制度

- 跨境数据流动政策成为国际政策热点。近期，欧盟法院裁决“美欧安全港框架”失效，美欧跨境数据流动政策分歧加剧。
- 我国应推动建立我国跨境数据流动政策：对跨境数据流动进行分级分类管理：涉及国家秘密、国家安全的数据严格禁止跨境，关键网络服务运营者应当本地存储用户个人数据等重要数据等。



厚德實學 興業致遠

---

政府高端专业智库  
行业创新发展平台

