

# 数据驱动 环境管理转型与创新

环境保护部信息中心

徐富春

# 主要内容

**数据的再认识**

**大数据应用实例**

**我们的思考**



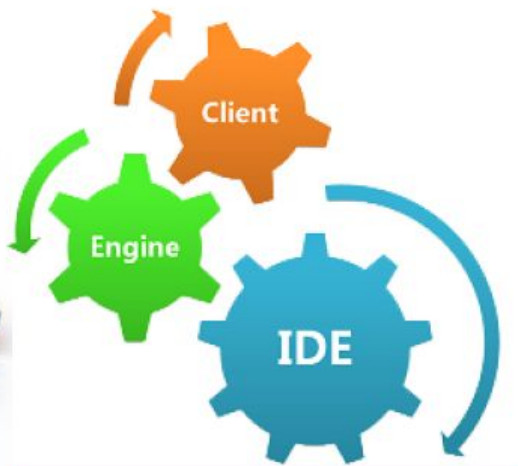
# 传统信息化时代

## 信息技术 Information Technology

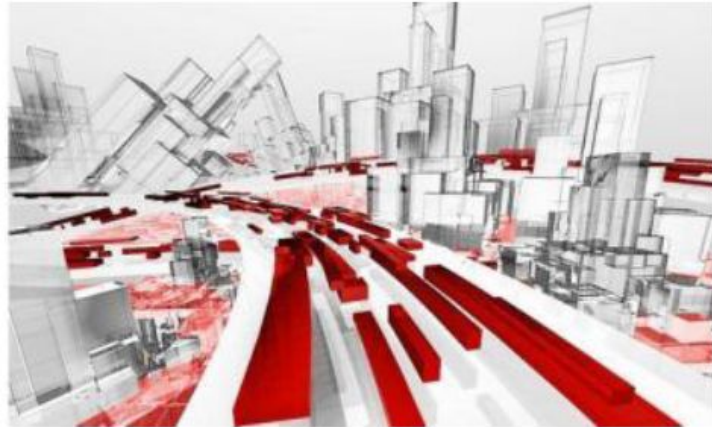


✓ 从286到386到双核，追求CPU处理速度

✓ 从C/S到B/S，从管理信息系统到辅助决策系统，追求应用系统的实现方式和功能



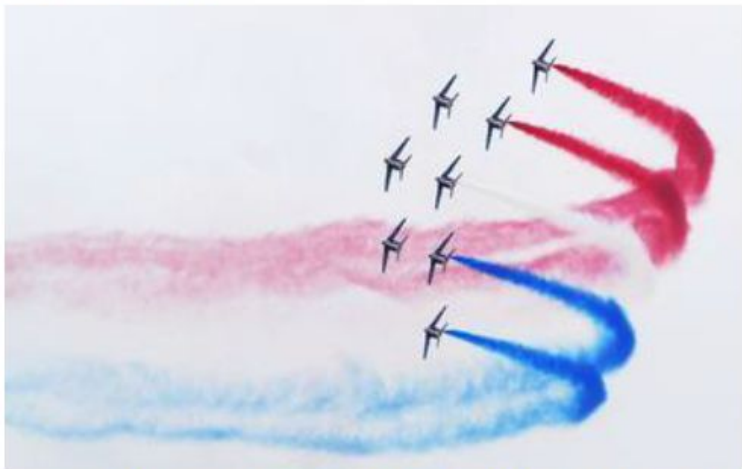
# 新技术的兴起



智慧城市、大数据、移动互联、云计算、物联网、社交媒体等新兴信息技术和应用，说明人类社会正在进入信息时代新阶段——以数据为主要资源，以网络化和智能化为主要目标的信息时代。

# 世界正进入ZB级数据计算时代

- ✓ **摩尔定律**：芯片性能每18个月倍增
- ✓ **数据摩尔定律**：从现在起，每18个月，新增的存储量等于有史以来存储量之和。



## 全世界信息技术能力发展的历史趋势 (1986-2030)

- ✓ 大体上每15-20年存储-通信-计算的能力增加1000倍
- ✓ 人均能力: Mega (兆,  $10^6$ ) → Giga (吉,  $10^9$ ) → Tera (太,  $10^{12}$ )
- ✓ 全球能力: Peta (拍,  $10^{15}$ ) → Exa (艾,  $10^{18}$ ) → **Zetta (泽,  $10^{21}$ )**

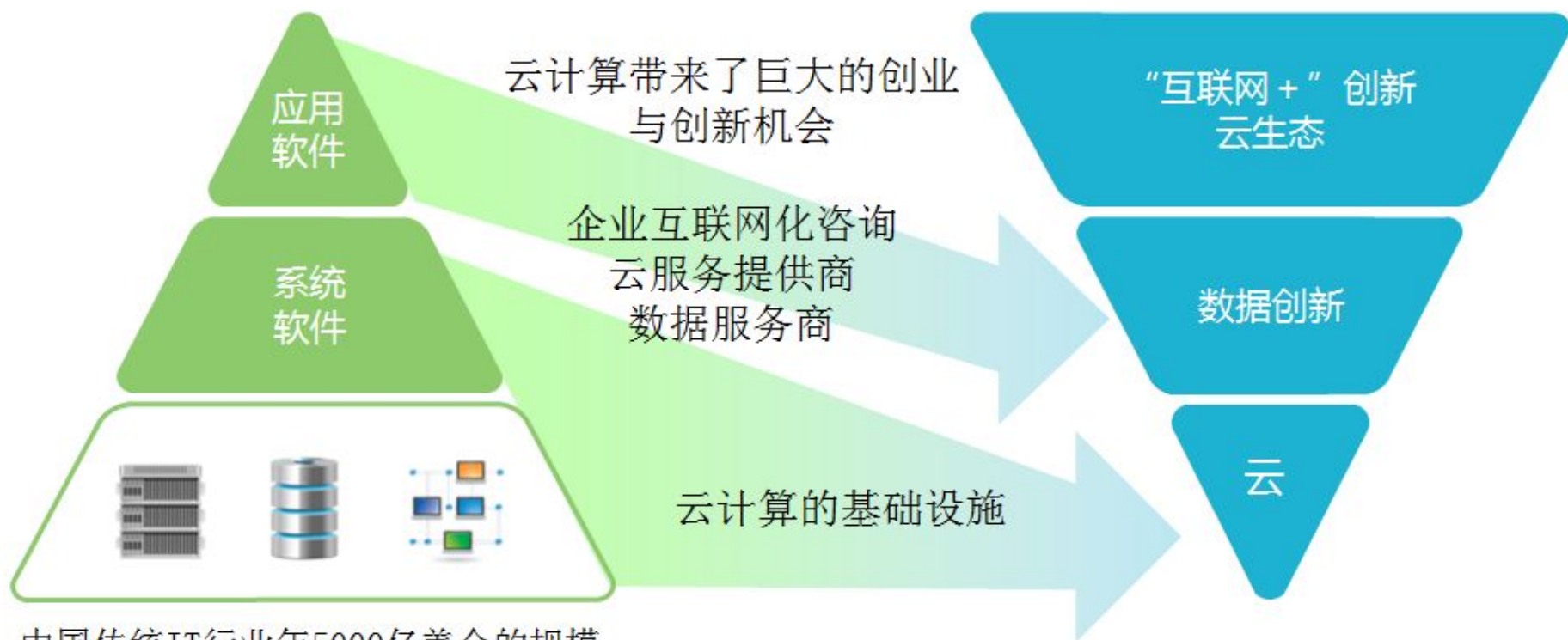
# 数据暴雨



2003年微软研究院的Tony Hey院士提出了“数据暴雨”的概念，2009年在《科学》杂志上又做了更新，指出我们才用了一点点数据的巨大潜能。

# 云计算与大数据支持“创新生态”

改变20年“重硬轻软”的惯性；计算机+软件 -> 云服务+数据



中国传统IT行业年5000亿美金的规模

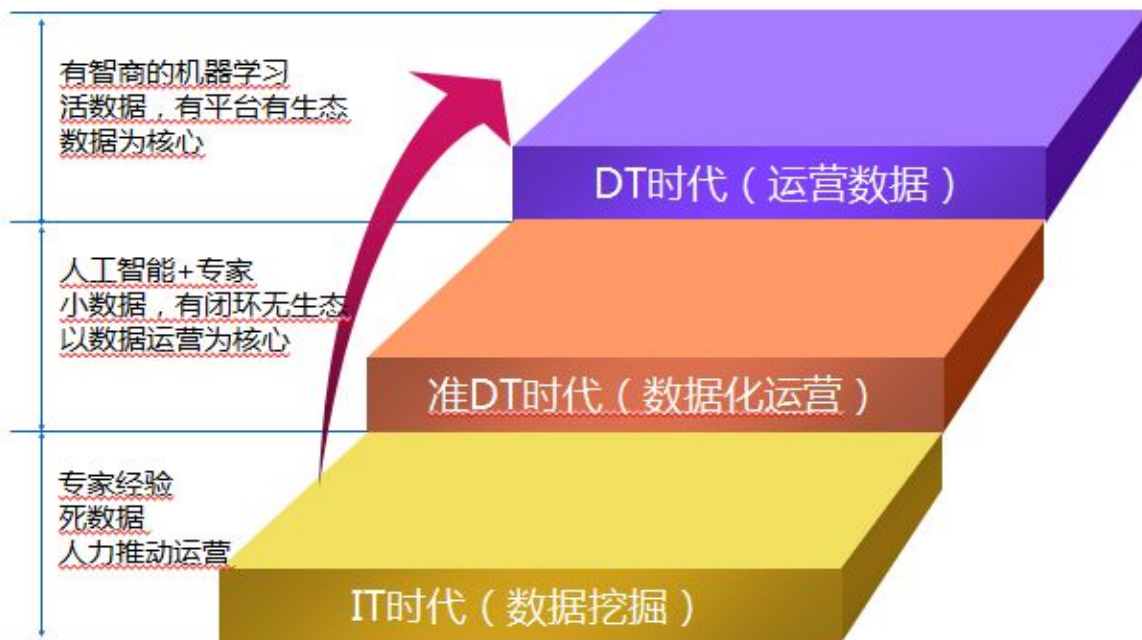
# 云计算与大数据全球趋势

45% vs 4%

云服务增长率：45%  
传统 IT 增长率：4%

云计算的本质是“资源聚合，提高资源利用率”

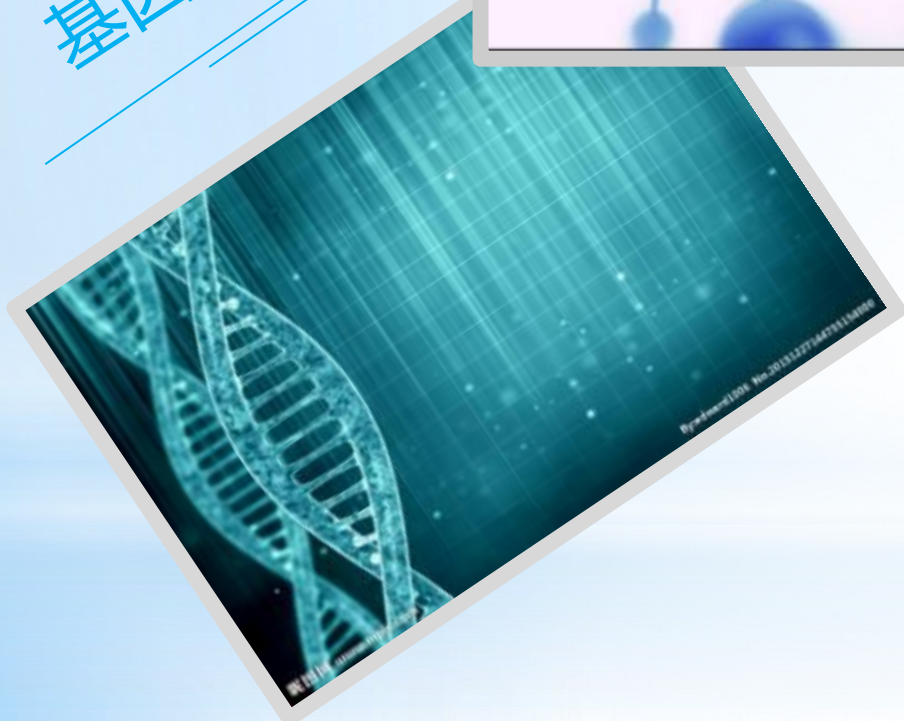
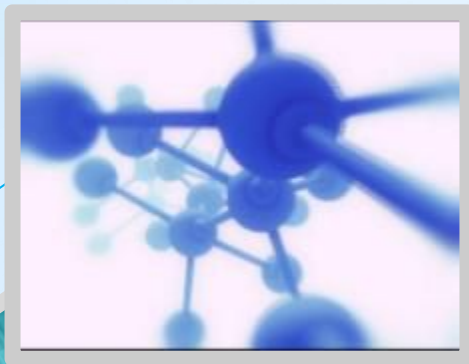
云计算是基础设施  
泛互联网化  
进入DT时代





# 数据的基因特质

基因特质



**数据具备基因特质**：可以复制、继承、分裂、重组、突变等。多种数据的**交叉关联**就像基因的交叉，可以衍生出混血儿、新物种，但也可能产生**变异**。数据已经演化成为生产要素，会对原有的生产要素配置产生冲击，形成**新的生产方式和方法**。

# 数据技术 Data Technology

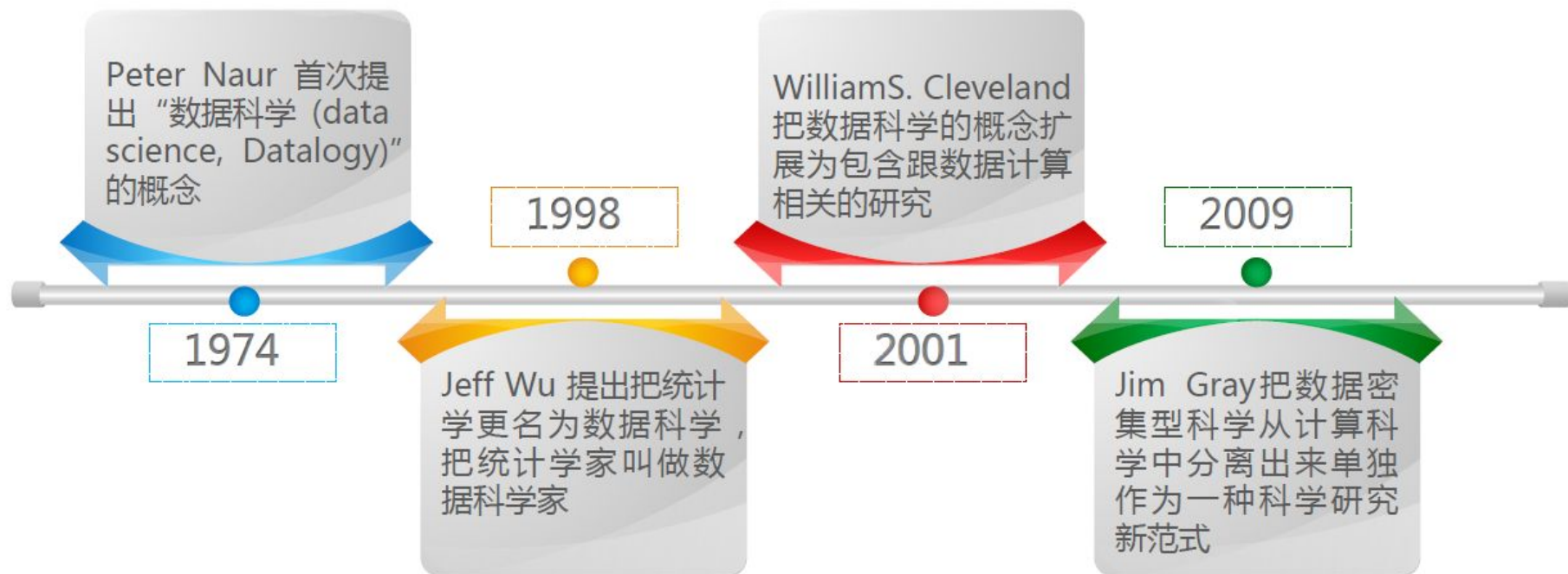
## (1) 核心问题：

如何从庞大复杂多样及变化的数据中经过处理分析得出有价值的管理和决策依据

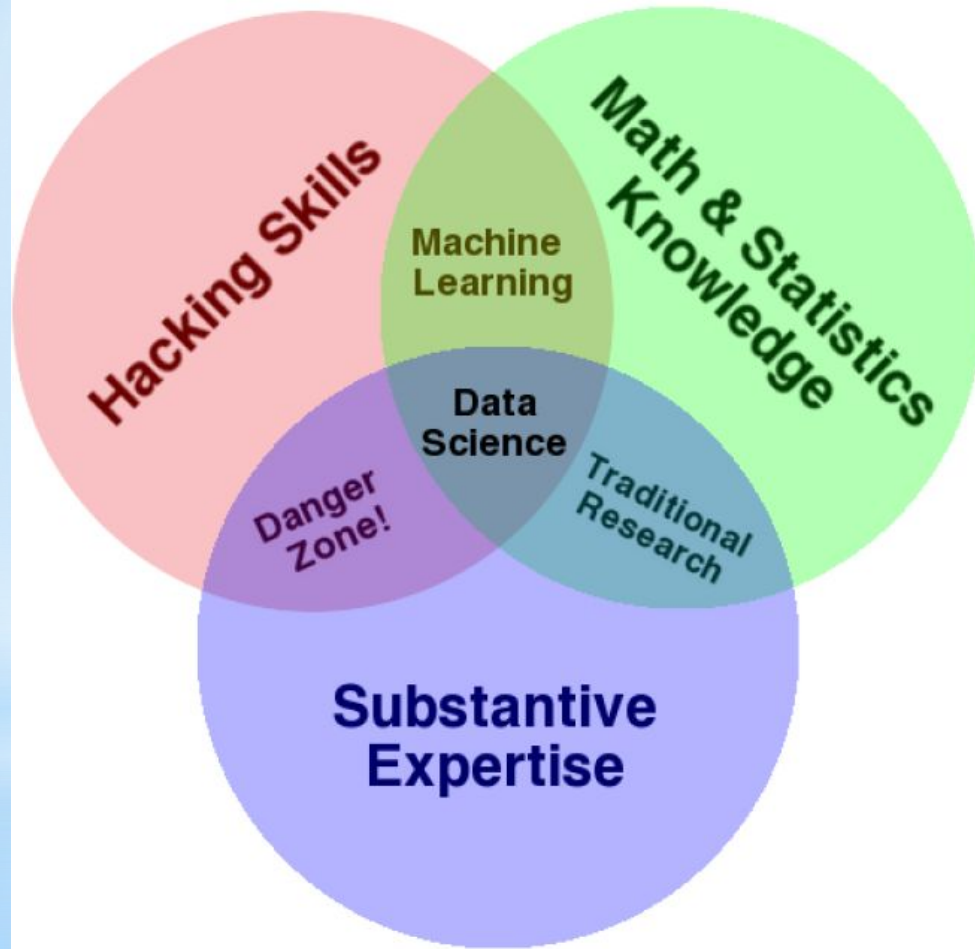
## (2) 数据科学 ( Data Science )：

数据中蕴藏巨大的价值，能够对数据进行分析、处理，挖掘出信息和知识，并对结果进行展示和解释，用于指导决策的科学

# 数据科学发展简介



# 数据科学的构成



计算机技巧——数据的获取、整理、运用  
Hacking Skills

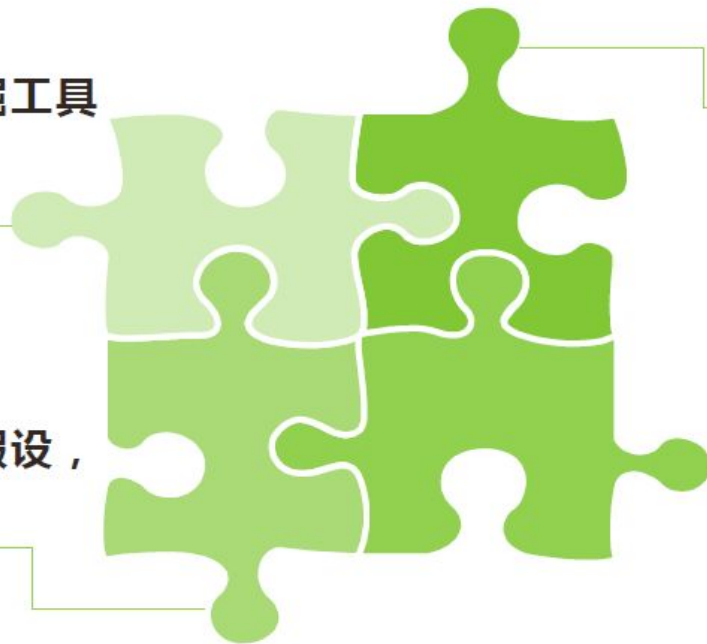
数学和统计知识——数据的挖掘  
Math & Statistics Knowledge

实质性的专业知识  
Substantive Expertise

# 数据科学解决问题的过程

用科学的方法、数据挖掘工具  
洞察新数据

充满好奇心，质疑现有假设，  
在数据中发现趋势



用探索数据的方式看待周围的世界

能够管理和洞察数据

# 数据处理分析的基本方法

## 数据分析

描述性分析

分析过去

预测性分析

预测未来

指令性分析

对行为的指导

认知性分析

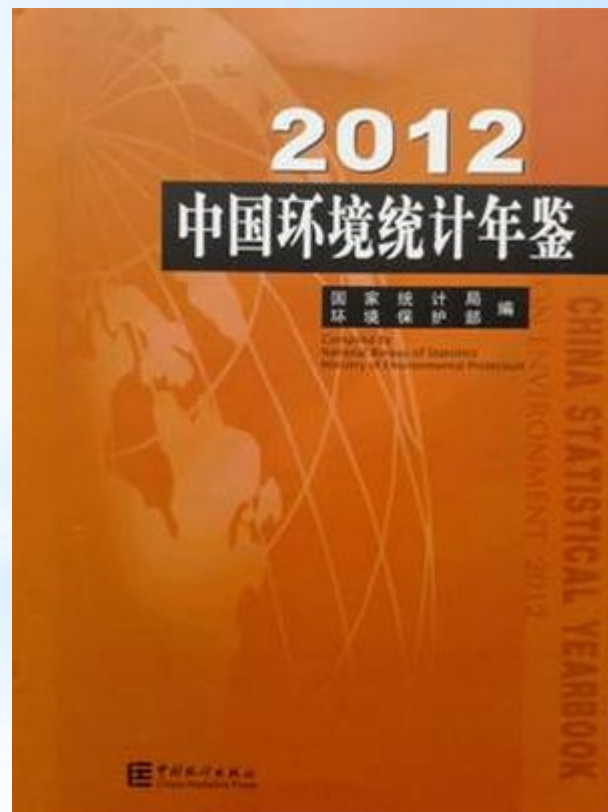
集大成者

- ✓ 描述性分析占数据分析的 90%
- ✓ 预测性分析占数据分析的 5%
- ✓ 剩下5%为指令性分析和认知性分析

# 数据分析——描述性分析

- ✓ 基于历史数据的滞后性分析，从数据出发概括数据特征，为业务措施提供清晰的评估。

一堆数据生成一个报表，对数据进行评估，约占90%

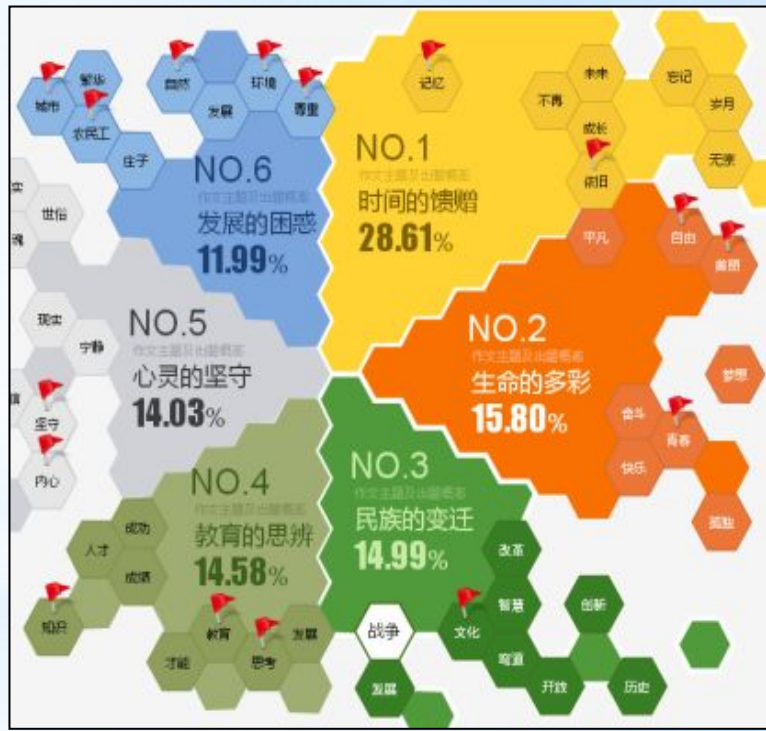


# 数据分析——预测性分析

- ✓ 预测分析是利用各种统计、建模、数据挖掘工具对最近的数据和历史数据进行研究，从而对未来进行预测。



依据一个算法  
通过挖掘和统计对数据走向  
和发展做出一个判断和预测，  
对数据作出一个判断，  
约占5%





# 数据分析——指令性分析

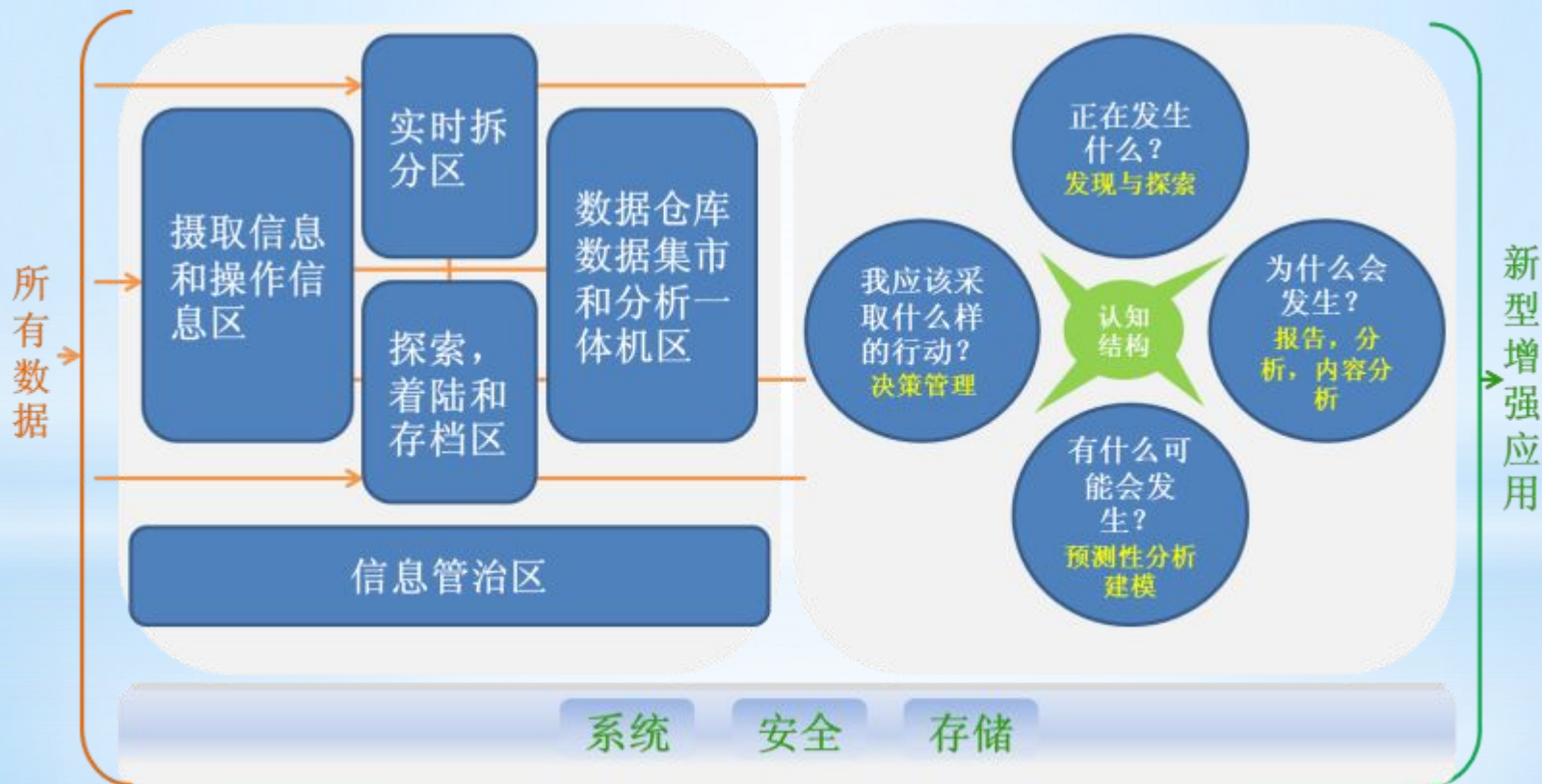
- ✓ 指令性分析是在分析过去和预测未来的基础上对行为的指导，分析每个环节、每个步骤、每个流程、每个岗位，信息更具准确性，业务决策者可以直接使用，**处于增长势头**。



预测性分析和指令性分析的区别在于，预测性分析我只告诉你，这只股票有可能会往下走，大盘有可能会跌，但它不是个指令，指令型分析的结果是说，当某某股票到21块5的时候进，到22块6的时候出，这叫指令型分析的结果。

# 数据分析——认知性分析

认知性分析是描述性分析、预测性分析、指令性分析的集大成者，借助于计算机自我学习和认知，提高对数据的自我处理和分析能力，提升数据处理极限和能力，对数据进行认知，形成数据思维能力。他不仅告诉你发生了什么、还会发生什么、应该采取什么样的行动，还告诉你为什么？



# DT的提出

2014年3月，马云在一次演讲中首次提到“DT”概念，指出“**人类正从IT时代走向DT时代**”。

**IT时代**是以自我控制、自我管理为主，而**DT时代**，是以服务大众、激发生产力为主的技术。这两者看似是一种技术的差异，但实际上是思想观念层面的差异。



# “第三次打车大战”



## 烧钱为人民服务？

通过移动支付搜集到消费者更多的行为数据，运用这些数据可以分析出消费者的爱好、需求、支付能力等等，由此更加精准地投放广告、获取市场信息、预测行业发展趋势……这些都是商家最需要得到的。

滴滴打车最终是要做智能交通，目标是做成一个数字化的交通平台，让大家的出行变得更方便。



共青团中央主办

充满智慧与创新精神的“阳谋”。

“滴滴”与“快的”用打车软件“芝麻开门”，积累大数据，获取大数据，为日后的开发大数据、利用大数据开辟“钱程”。

# “小米手机”



- ✓ 2014年三季度，小米智能手机出货**1840万台**；
- ✓ 2014年双十一活动中，小米在天猫平台上销售手机**116万台**，销售额**15.6亿元**，约占天猫当天总额的3%，成功卫冕单店第一；
- ✓ 2014年10月1号，小米云仅在中国**每天**存储的数据已经到了**380T**，明年对今年，最少还有**三倍**的成长。

2014年12月4日接受采访时，雷军表示：

“所以如果我们不能用**Big Data**的技术，转化出价值，那我们公司再撑下去就真的破产了……但是我扛得住，明年我也扛得住，后年我也扛得住，大后年Big Data要没价值的话，那我就破产了。”



# 数据技术和云计算

数据应用是关键，云计算和平台是大数据的基础和必由之路

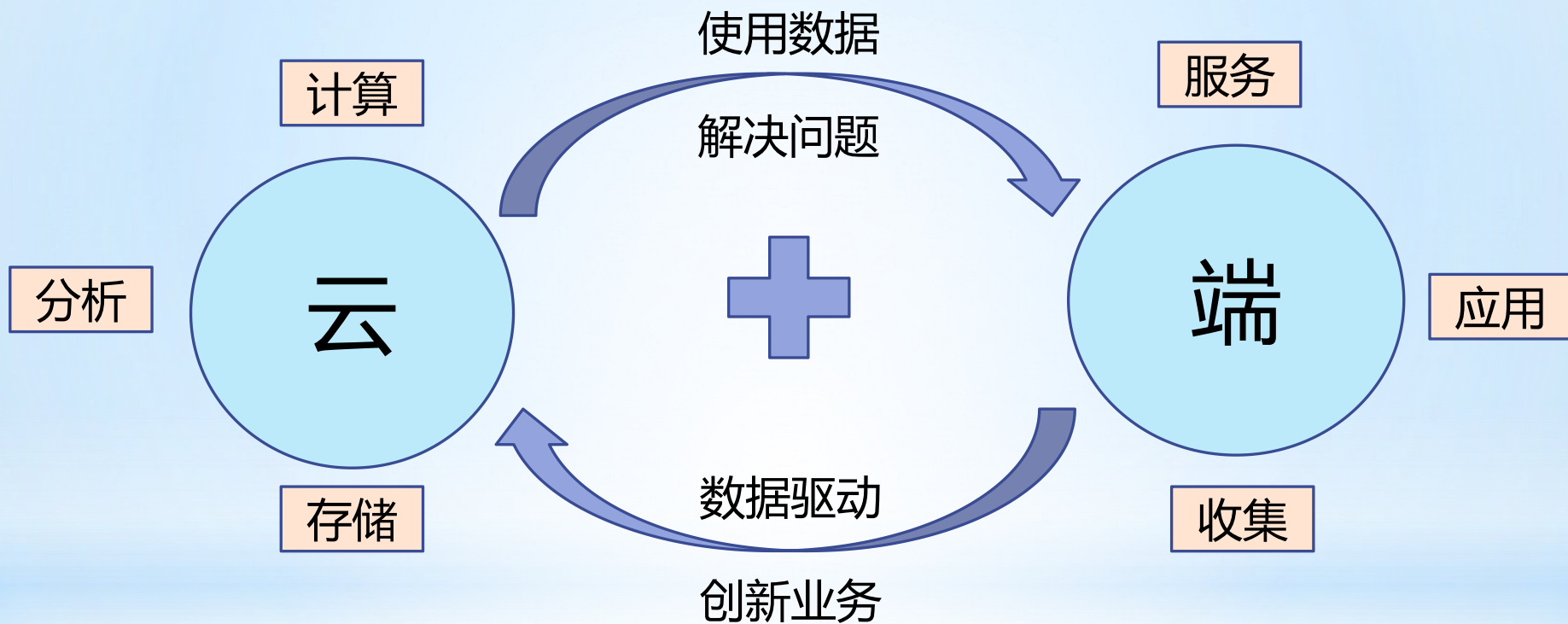
如果说数据是财富，那么大数据就是宝藏，而云计算就是挖掘宝藏的利器！



没有强大的计算能力，数据宝藏终究是镜中花；没有数据的沉淀，云计算也只能是杀鸡用的宰牛刀！



# 数据驱动信息化架构模型



# 几点认识

- \* 数据独立性：数据与应用分离，提高DBMS的适应性和应用系统稳定性。
- \* 关系型数据库：关系模型极大地提高了数据库管理能力，并形成了生态系统。
- \* Hadoop：丢失了大多数DBMS的特性，不支持事务或只支持简单事务，缺乏开发工具和环境，没有索引等，并非完美！
- \* 数据库核心技术是大数据管理和处理技术的基础。
- \* 大数据不是也不能取代传统的数据库技术。
- \* 大数据与其他数据处理技术相结合，通过细分的技术和市场，满足用户不同层面的需求



# 什么是大数据？

- \* 大数据（Big data），或称巨量资料、海量资料，指的是所涉及的资料量规模巨大到无法通过目前主流软件工具，在合理时间内达到获取、管理、处理、并整理成为帮助企业经营决策更积极目的 的资讯。
- \* 大数据（Big data）是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产。
- \* 大数据的定义

大数据是指无法在一定时间内用常规软件工具对其内容进行抓取、管理和处理的数据集合

大数据≠海量数据。用传统算法和数据库管理系统可以处理的海量数据不等于“大数据”

大数据=“海量数据”+“复杂类型”+“快速处理”

# 数据新思维

- \* 大数据时代使得数据具有独立存在的价值，数据作为**基础设施**的地位突显。
- \* 以前都是先有应用后有数据，现在是先有数据再说应用。
- \* 数据为业务服务，软件为数据服务。
- \* **建立以数据为核心的信息化新思维。**
- \* **要全部数据，不要抽样。**
- \* 模型不再那么重要，让数据说话，“All models are wrong, but some are useful.” “All models are wrong, and increasingly you can succeed without them.”

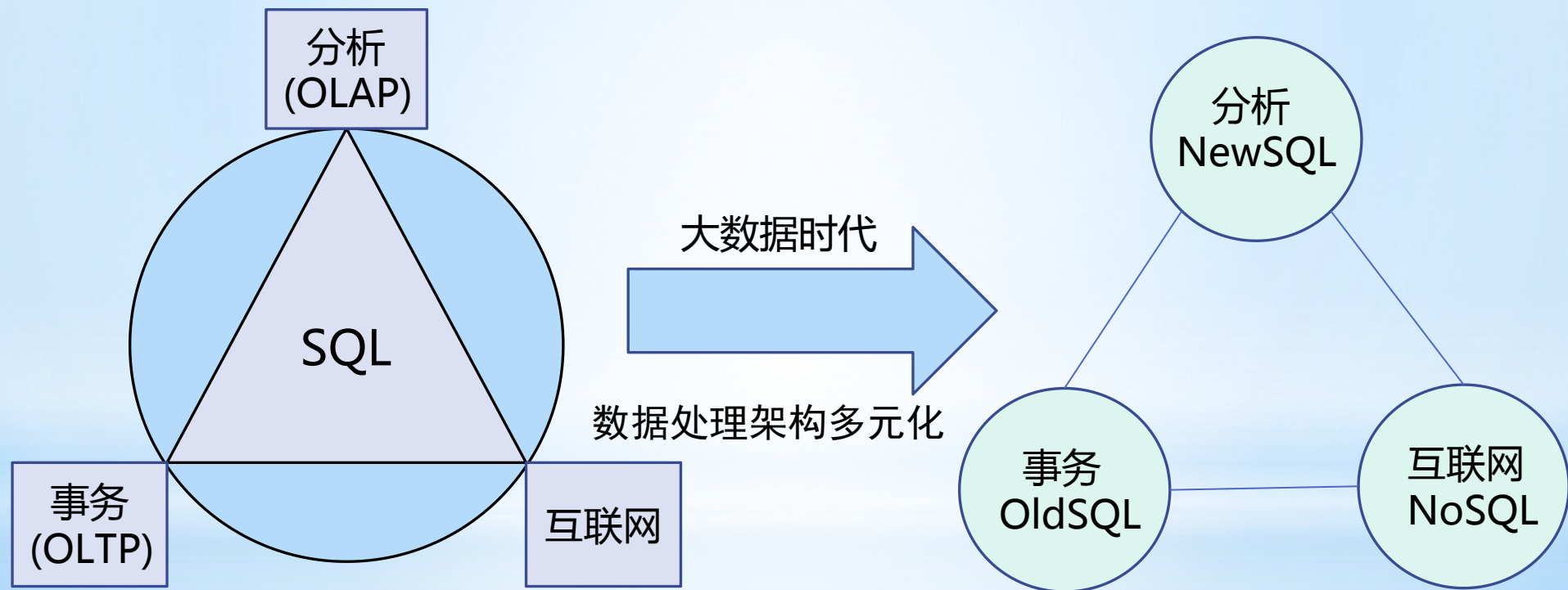
**数据库技术与大数据技术成为数据管理和数据处理领域  
两大重要的支撑技术**

# 大数据推动数据处理架构变革

多种架构支撑多类应用

( One Size Does Not Fit All ,  
一招鲜不灵了 )

( One Size Fits All , 一招鲜 )



# 主要内容

DT的再认识

大数据应用实例

我们的思考



# 环境形势严峻



水污染



大气污染



固体废弃物污染



酸雨



荒漠化



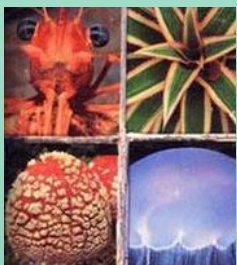
森林锐减



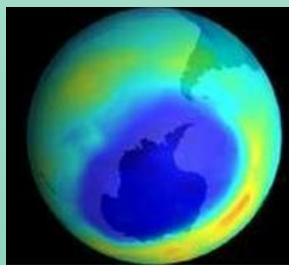
资源短缺



耕地减少



生物多样性丧失



臭氧层损耗



全球气候变化



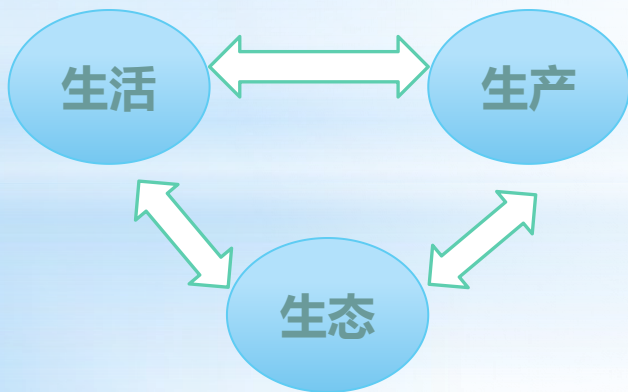
持久性有机物污染

# 十八大提出建设“美丽中国”



建设生态文明，是关系人民福祉、关乎民族未来的长远大计。面对资源约束趋紧、环境污染严重、生态系统退化的严峻形势，必须树立**尊重自然、顺应自然、保护自然的生态文明理念**，把**生态文明建设放在突出地位**，融入经济建设、政治建设、文化建设、社会建设各方面和全过程，努力建设**美丽中国**，实现中华民族永续发展。

## 对“美丽中国”的理解



**“三生” 共赢**

是文明发展的高级阶段

吸取中国古代生态智慧“天人合一思想”

青山、绿水、蓝天，是我们留给子孙后代最宝贵的财富，是时代赋予我们的庄严历史使命



## 宏观环境

中共十八大提出：

1. 提升城镇化水平，推动城乡发展一体化——城镇化与城乡发展一体化；
2. 重点抓好四个方面的工作：一是要优化国土空间开发格局；二是要全面促进资源节约；三是要加大自然生态系统和环境保护力度；四是要加强生态文明制度建设。

用创新思维推进环保战略转型：

从单纯的环境保护到生态环境，再到生态文明。  
从广义上讲，人民群众是环境保护的重要主体，从政府包办环保向发挥政府、市场、社会综合作用转变。

# 大气污染防治行动计划

🌍 又称“**大气十条**”，2013年6月，由国务院发布。确定了**十项具体措施**，是当前和今后一个时期全国大气污染防治工作的行动指南。



🌍 具体目标：

- 到**2017**年，全国地级及以上城市可吸入颗粒物浓度比2012年下降10%以上，优良天数逐年提高；
- 京津冀、长三角、珠三角等区域细颗粒物浓度分别下降25%、20%、15%左右；
- 其中北京市细颗粒物年均浓度控制在60微克/立方米左右。



# 水污染防治行动计划

🌍 又称“**水十条**”，2015年4月，由国务院发布。将在污水处理、工业废水、全面控制污染物排放等多方面进行强力监管并启动严格问责制，铁腕治污将进入“新常态”。



🌍 工作目标：

- ▶ 到**2020**年，七大重点流域水质优良（达到或优于Ⅲ类）比例总体达到70%以上，地级及以上城市黑臭水体均控制在10%，地级及以上城市集中式饮用水水源水质优良比例总体高于93%，近岸海域水质优良比例达到70%左右；
- ▶ 到**2030**年，力争全国水环境质量总体改善，水生态系统功能初步恢复；
- ▶ 到**本世纪中叶**，生态环境质量全面改善，生态系统实现良性循环。

# 环境保护与大数据

\***环境保护**面临新形势，最重要的是如何转变思想，有效积极应对，找出有效途径

\***大数据**开启了一次重大的时代转型，大数据正在改变我们的生活及理解世界的方式，成为新发明和新服务的源泉，而更多的改变正蓄势待发.....



# 互联网企业引领全球大数据应用

互联网和移动  
互联网快速产生  
的各类数据

用户行为数据

网页数据

系统日志数据

用户交易数据

.....

## 互联网企业大数据应用场景

### □定向广告

Google

facebook

YAHOO!

▶如Facebook对用户基本属性、行为习惯和兴趣等进行语义分析，为广告商提供基于数据挖掘的自助式广告下单服务系统

### □个性化推荐

amazon.com

淘宝网  
Taobao.com

▶如亚马逊利用大数据技术为用户提供社会化推荐、广播式个性化推荐等服务，加快了产品传播的速度

### □满意度分析

twitter

Apple

▶如Twitter对提到产品的文本进行搜集并按规则打分，得到客户对产品的满意度评价

### □提升服务质量

Google

facebook

▶如Facebook对大量用户产品使用状况的数据进行分析，优化产品设计及服务，改善用户的使用体验

### □社会服务

Google

Baidu 百度

▶如谷歌基于用户搜索数据推出的产品谷歌流感趋势，可以迅速、准确的对流感进行预报

### □其它

▶.....

互联网企业  
大数据应用  
基本特点

□定向广告和  
个性化推荐

□简单的大数  
据应用已在互  
联网领域广泛  
开展，且大部  
分企业具备自  
行实施应用的  
技术能力

□掌握有大量  
用户行为数据  
的互联网巨头  
可以较好提供  
社会化服务

# 阿里大数据：超越Hadoop的大规模集群计算平台

## 自主可控

- 数据仓库，数据挖掘
- 数据分析，数据分享

## 海量数据

- 海量结构化，半结构化
- 实际容量500+PB

## 海量计算

- 100T Terasort < 30Min
- 生产集群 > 20000
- 20w Job, 日处理10PB

## 计算模式

- K-V, MPP
- Hive SQL, Map Reduce
- Graph, MPI
- 流计算

## 机器学习平台

- 机器学习算法
- MPI, BSP, Graph, PS

## 专业团队

- CDO
- iDST
- 安全, 蚂蚁金服...



# 阿里大数据工作台与市场

## 应用场景

DW

数据仓库 & BI  
海量数据加工处理



数据挖掘  
海量数据挖掘



数据开放  
多方数据交易



数据业务应用  
智能推荐 | ...

大数据平台  
BASE

交换平台

数据市场

应用市场

数据服务化接口：实时 | 离线

开发平台

开发  
套件

BI  
套件

数据探  
索环境

数据  
中心

运维  
平台

数据同步：实时 | 离线

计算平台

离线计算引擎

实时OLAP引擎

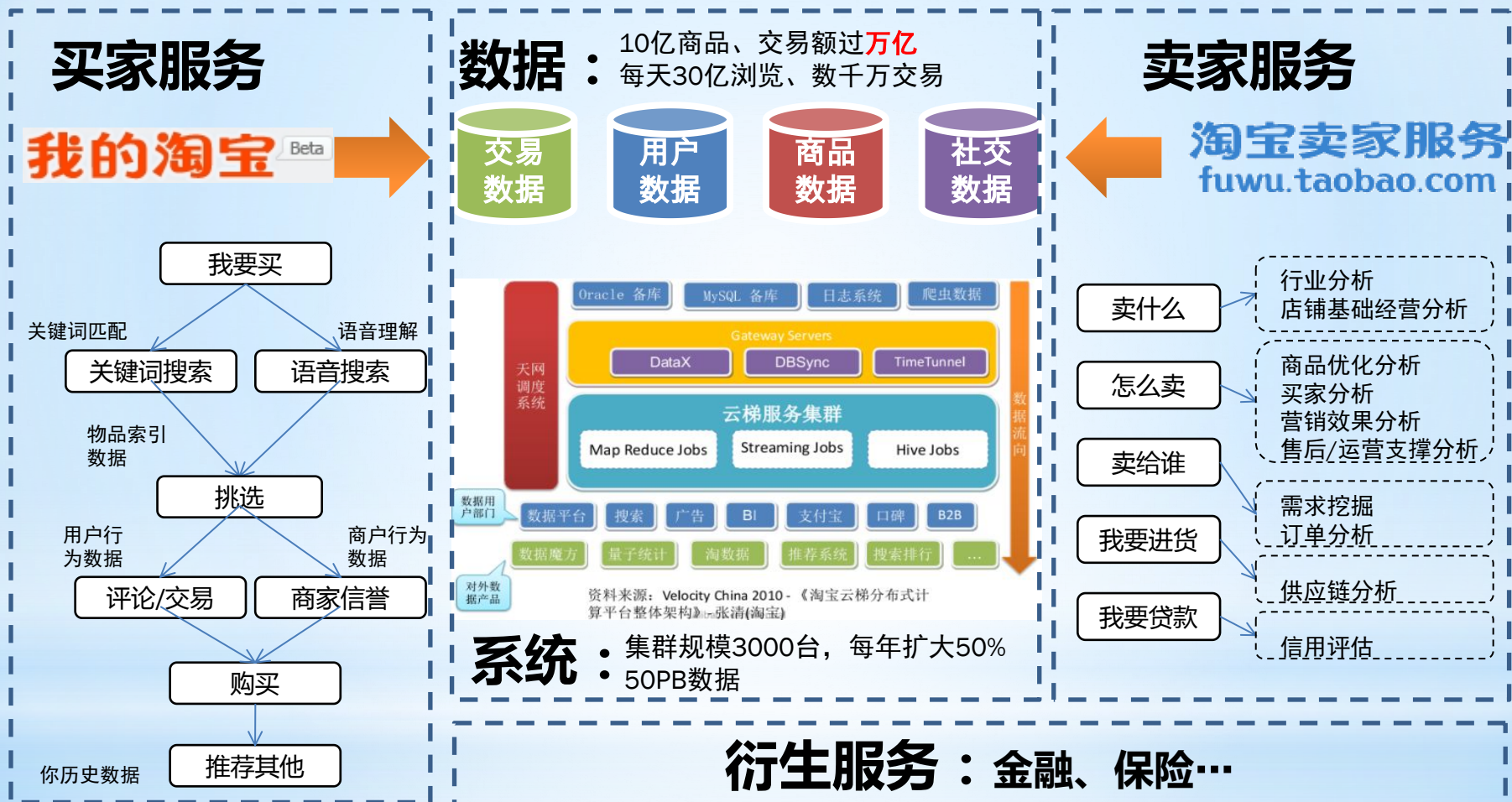
流式计算引擎

统一  
一元数据服务  
统一  
账号体系



阿里云基础设施

# 淘宝案例



**搜索、电商、广告、SNS等数据驱动的互联网服务取得巨大成功，激发了大数据应用的想象力！**

# 其它行业普遍仍在探索...

传统行业

## 医疗

➢美国DNAnexus为医疗机构和用户提供了基因数据的管理、分析和可视化能力



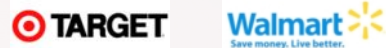
## 能源

➢能源机构Vestas综合考虑温度、降水、风速、湿度、气压等因素，确定风力涡轮机的最佳安置地。



## 零售业

➢沃尔玛零售数据商业智能分析系统，可以了解到全球4000多家门店每天的销售情况并辅助制定销售策略。



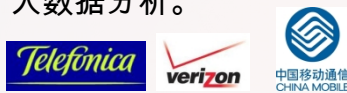
## 制造业

➢日本小松公司根据挖掘机工作情况进行大数据分析，判断下一年度的市场需求。



## 电信运营

➢西班牙电信“智慧足迹”产品可提供基于位置的大数据分析。



## 金融

➢美国征信所Equifax对海量信息进行交叉分析，推出70余项新服务。



其它：农业、气象.....

## 传统行业大数据应用基本特点

- ❑数据源主要来自企业内部、类型较少、实时要求较低
- ❑企业逐渐重视大数据，但当前应用相对简单，处于探索阶段
- ❑掌握大数据技术的企业较少，主要由IT企业提供技术支持

## 互联网与传统产业不断融合，将催生新的大数据创新机会

互联网与传统行业融合创新大数据应用和产品

### 金融与互联网融合

➢阿里小贷：基于对用户交易行为的大数据分析，为阿里面向中小企业实施信用贷款提供支撑



### 交通与互联网融合

➢德国电信利用大数据技术实施德国政府的无拥塞交通研究项目



其它：  
.....

❑融合发展能够将互联网的在线、数据快速积累和获取等优势带至传统行业，为实体经济发展带来新的突破，将是未来大数据发展的重要方向。

# 环保的棘手问题可以用大数据来解决吗？



**环境保护**

——还土地以绿色



**大数据**

——应用大数据技术





# 大数据技术应用于环保已具备先决条件

## A

### Attention

- 社会各界在环境保护上有共同的价值关注点

## B

### Blast

- 大数据的存储、挖掘和应用技术有了爆炸式发展

## C

### Complicated

- 环境保护的对象、利益相关人、措施等之间关系复杂，可挖掘大量有价值的联系

## D

### Data

- 功能丰富、采集数据准确的微小传感器正在积累大量数据

# 大数据技术应用于环保具有发展潜力

## 时间维度

- **过去**：环境污染事故的大数据用于构建预测模型的训练
- **现在**：环境状态变量的大数据与出行、疾病预测、城市资源配置联系起来，提供城市美化解决方案
- **未来**：以海陆空各类环境变量构建统一视图和大数据应用开发平台，在最小伤害的情况下充分利用环境资源，维持可持续性发展

## 行业维度

- **医疗**：通过研究城市的部分环境变量与疾病的关系，建立医疗行业的易患疾病模型，开发对应医疗产品
- **家电零售**：通过分析家电使用场景，构建智能节电的家电研发模型，降低家庭的能源消耗，提升家电零售的环保质量
- **化工**：通过研究化工厂选址对环境变量的影响，对化工厂选址的决策提供支持；研究工厂排污单元的管控，降低排污成本

## 主体对象维度

- **政府**：基于污染指标的城市环境监测，基于气象、地质监测的灾难提前预警，基于环境疾病模型的疾病预警和监控，基于位置信息的生活垃圾智能化处理...
- **企业**：基于排污处理管控能力提升的成本控制，基于环境变化预测的医疗、健康产品开发，基于碳排放资源数据的分析管理能源结构、降低成本...

# 基于大数据的城市空气质量细粒度分析和预测（一）



气象



交通



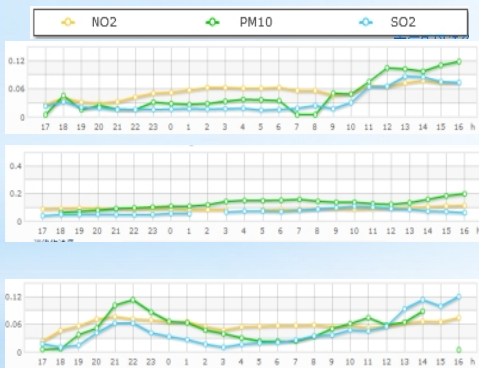
人类活动



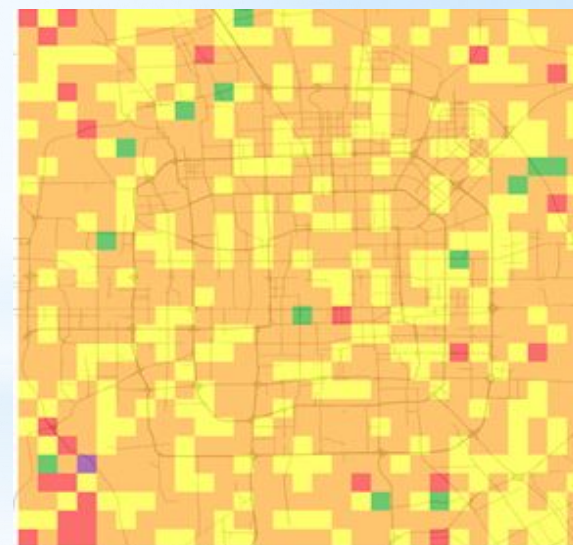
兴趣点 (POIs)



道路网

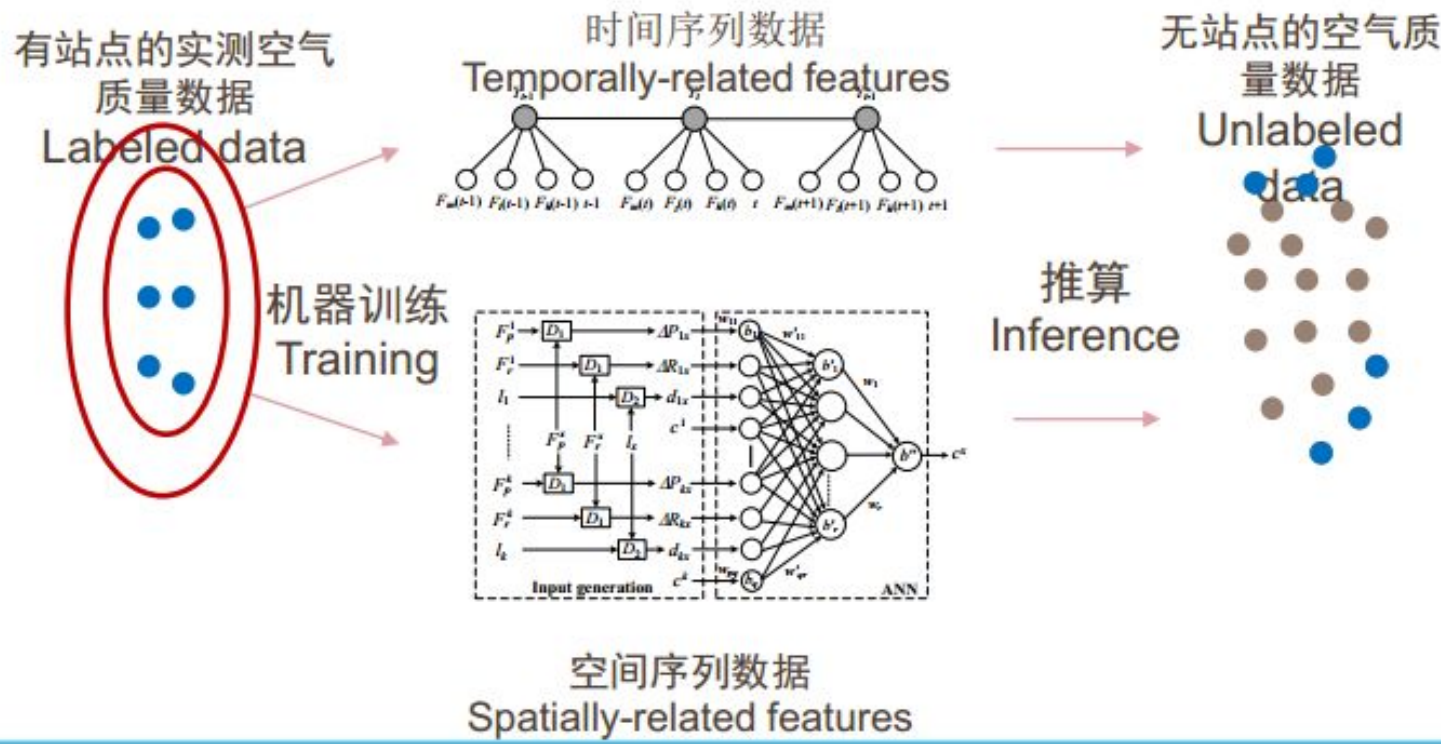
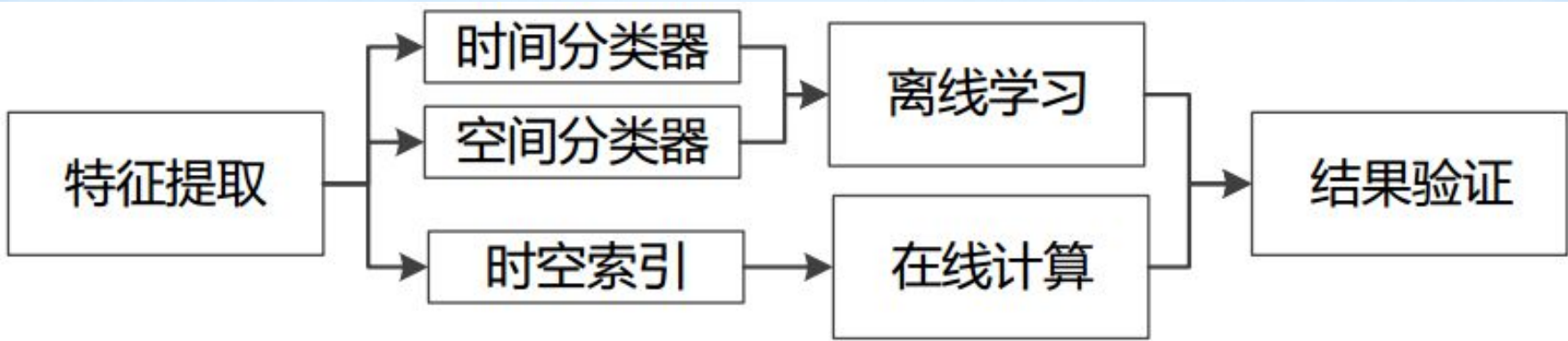


实时监测站空气质量报告

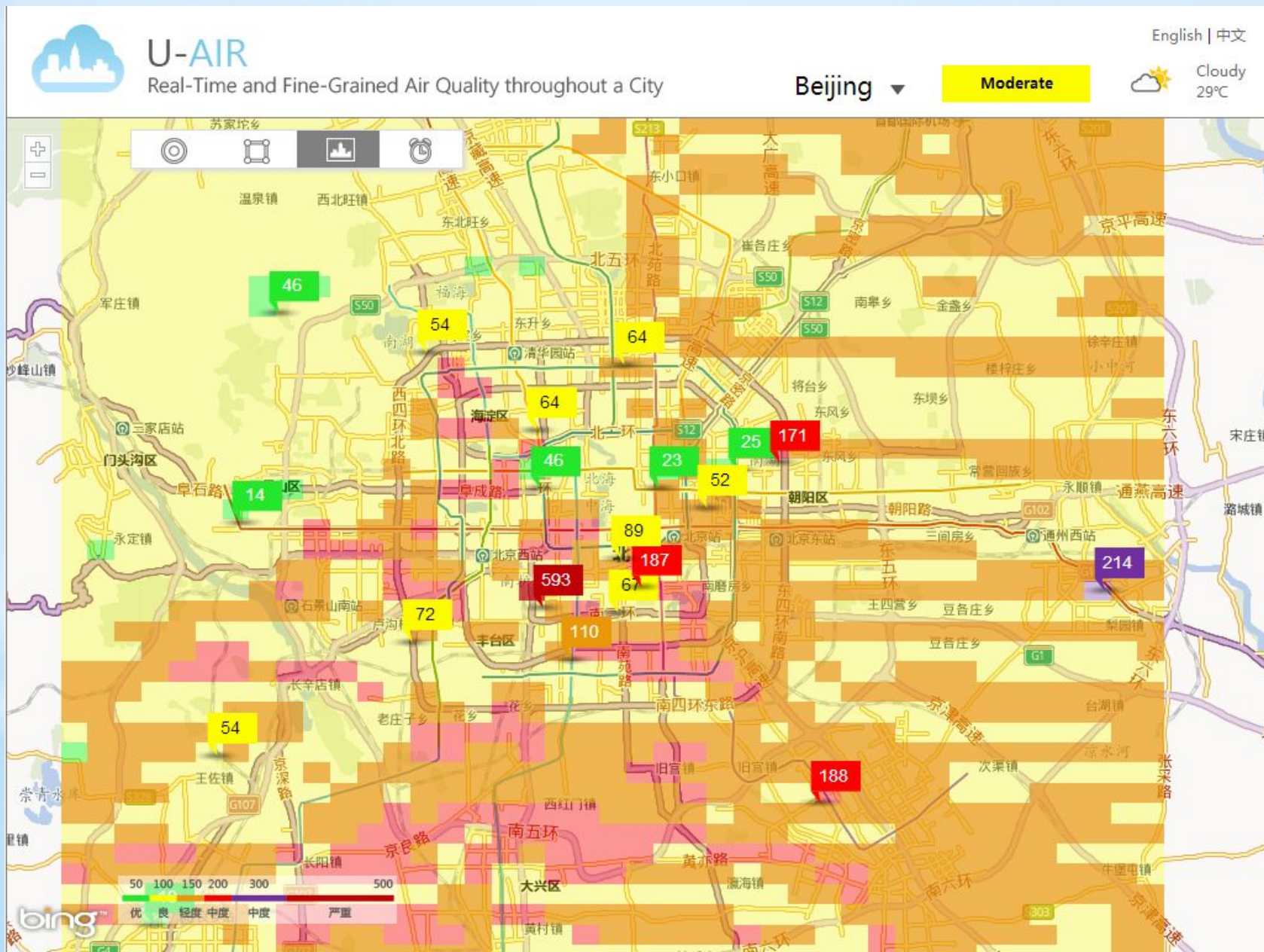


城市局部地区空气质量

# 基于大数据的城市空气质量细粒度分析和预测（二）



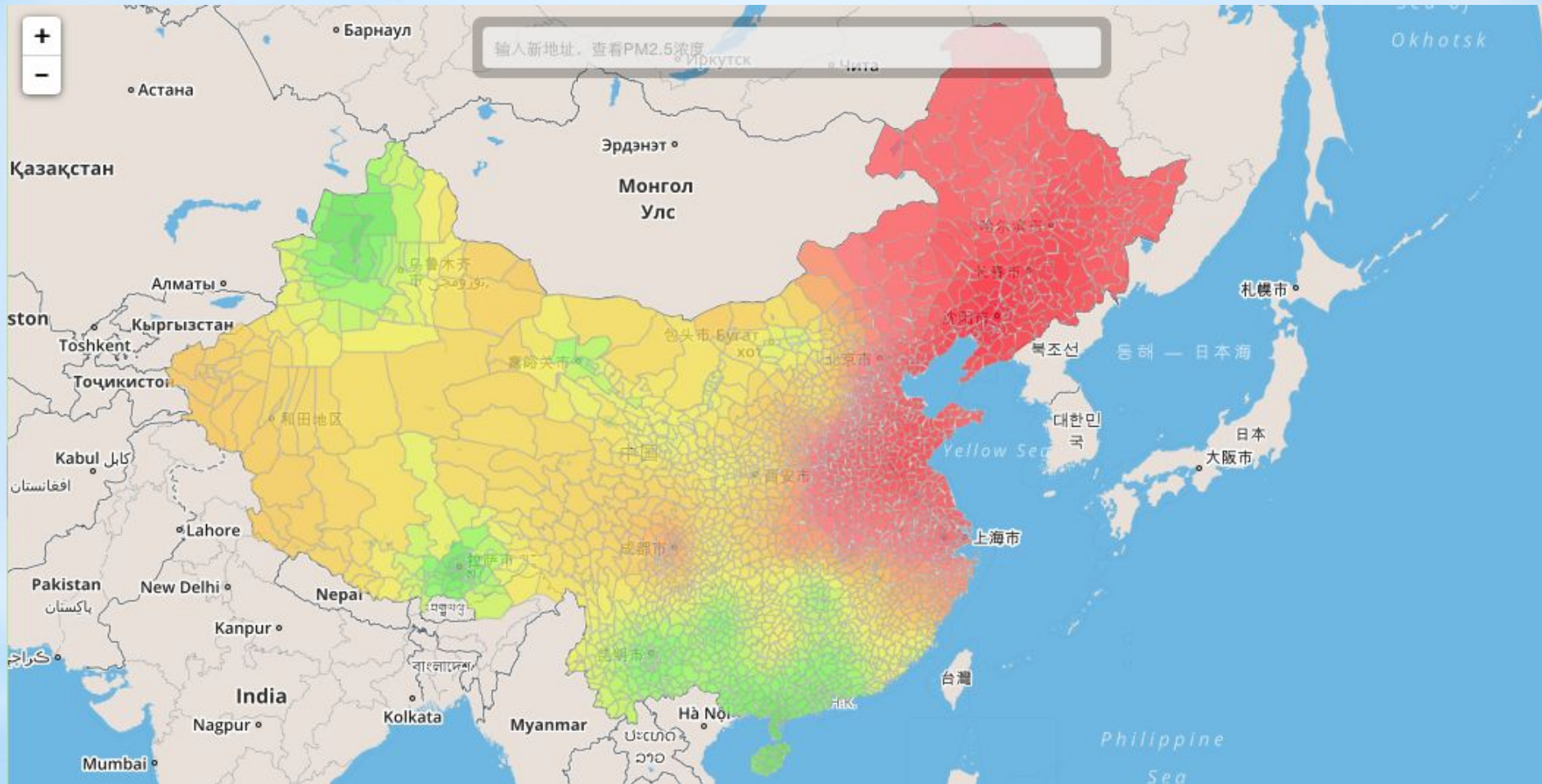
# 基于大数据的城市空气质量细粒度分析和预测（三）



# 利用卫星数据预测空气质量（一）

- \* **卫星遥感技术**还可以通过对灰霾层厚度和范围的监测，然后与地面监测站点的**监测数据**融合，计算出高时间和空间精度的地表PM2.5的浓度
- \* 目前我国地面空气质量监测点位的设置是根据行政区域设置，**站点数量**非常有限，不能反映污染物的空间分布，易受站点附近环境影响，造成不同区域受众对数据可信性的质疑。
- \* 卫星遥感监测技术能完全弥补地面监测在地域上的局限性，做到每一寸土地都能实时监测到，完全没有空白。同时可以根据整个区域内的污染图像做出更客观的全区域性污染指数估算。

# 利用卫星数据预测空气质量（二）



佳格霾图（GAGO）：基于卫星遥感大数据，与地面监测站点数据实时融合，采用机器学习算法，实时计算并显示全中国范围内的PM<sub>2.5</sub>浓度。每小时更新一次，空间精度为五公里

# 流域水污染控制应用案例 TMDL模型（一）



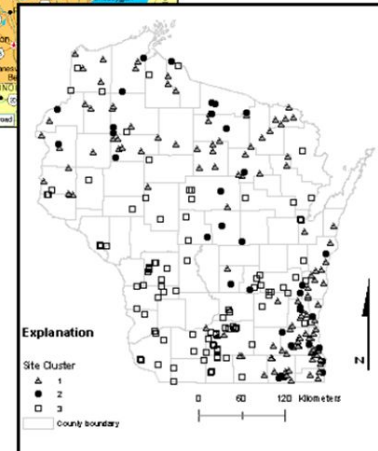
# 流域水污染控制应用案例 TMDL模型（二）

大数据模型用于流域地表水浓度预测、TMDL分配等，在很多情况下比传统水文、水质模型具有更好的模拟精度和适用性，应用于美国多个流域的水质管理中。

截至2011年,美国累计实施TMDL计划的总数量已达45000多个，改善水体质量的效果非常明显。

## Wisconsin streams temperature model

- **Area:** Entire state of Wisconsin
- **Need:** To predict stream temperature for stream segments throughout state for fisheries management
- **Data:**
  - Stream temperature – half-hourly signals from June 1 to August 31 in 254 streams from 1990 - 2002
  - Climate – 353 signals across state; 7 air pressure, 156 air temp., 13 dew point, 164 precipitation, 13 solar radiation
  - 42 categorical parameters to describe stream and basin attributes

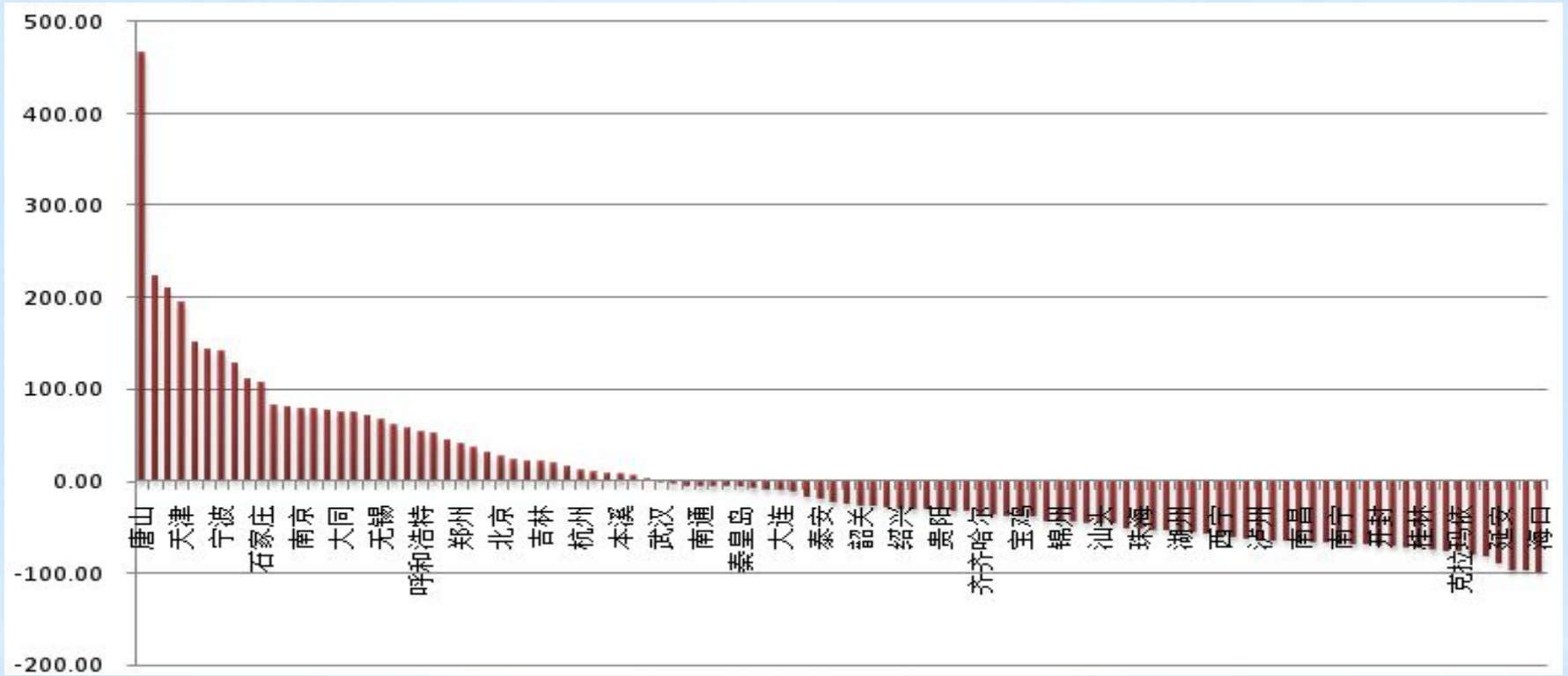


# 流域水污染控制应用案例 TMDL模型（三）



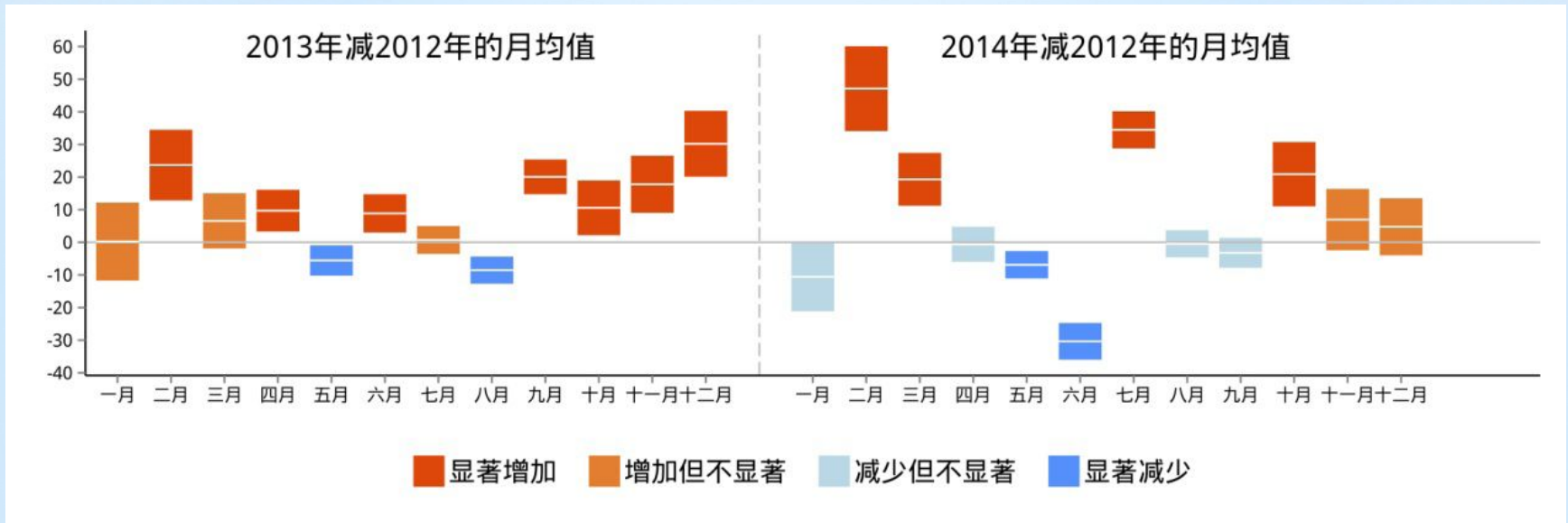
- ❖ 美国地质调查局（USGS）开发的大数据模型能够模拟点源的BOD、NH<sub>3</sub>排放和非点源的DO排放；
- ❖ 为该地区的排污许可证发放提供支持；

# 环境统计数据分析与应用（一）



以前仅仅通过分析**环境统计数据**计算各城市**综合污染指数**，没有考虑各城市在地理位置、气象因素、人口、经济等因素方面的差异，得出的结果不令人信服。

# 环境统计数据分析与应用（二）



🌐 剔除气象因素后，同2012年相比，2013年和2014年北京空气质量并没有改善，冬天的污染比2012年明显加重。

🌐 报告来自北京大学统计科学中心、北京大学光华管理学院


# 环境统计数据分析与应用（三）




# 污染源普查数据核查方法改进



# 环保数据废气和黑暗数据

 **数据废气 (Data Exhaust)**：数据价值和作用未充分利用和发现即被废置。很多**环保部门**都有数据废气，它的再利用价值也许现在不清楚，但在未来大数据的某一刻，它会迸发出来，可以化废为宝。

 **黑暗数据 (Dark Data)**：针对单一目标而收集的数据，一次使用即被闲置，其真正价值未能被充分挖掘。很多**环保业务**中存在黑暗数据现象。在大数据的情况下，黑暗数据可以用作其他用途，发挥更多效益。

 **结论：**

- 一方面环保决策和管理需要大量数据
- 另一方面环保数据又处于极大浪费中

# 主要内容

数据的再认识

大数据应用实例

我们的思考





# 数据驱动设想与效益

## 🌐 基本设想：

- 通过数据读懂环境变化与趋势（管理）
- 通过数据把握环境发展与动态（决策）
- 通过数据创新环境管理方法（改革）



## 🌐 效益分析：

- 数据驱动决策，提高预测准确率，决策更科学，更准确
- 数据驱动流程，把握业务流程更准确，提高业务工作效率
- 数据驱动产品，产品设计更个性化，服务更好更优
- 数据驱动管理，管理创新能力更强，转型升级更快

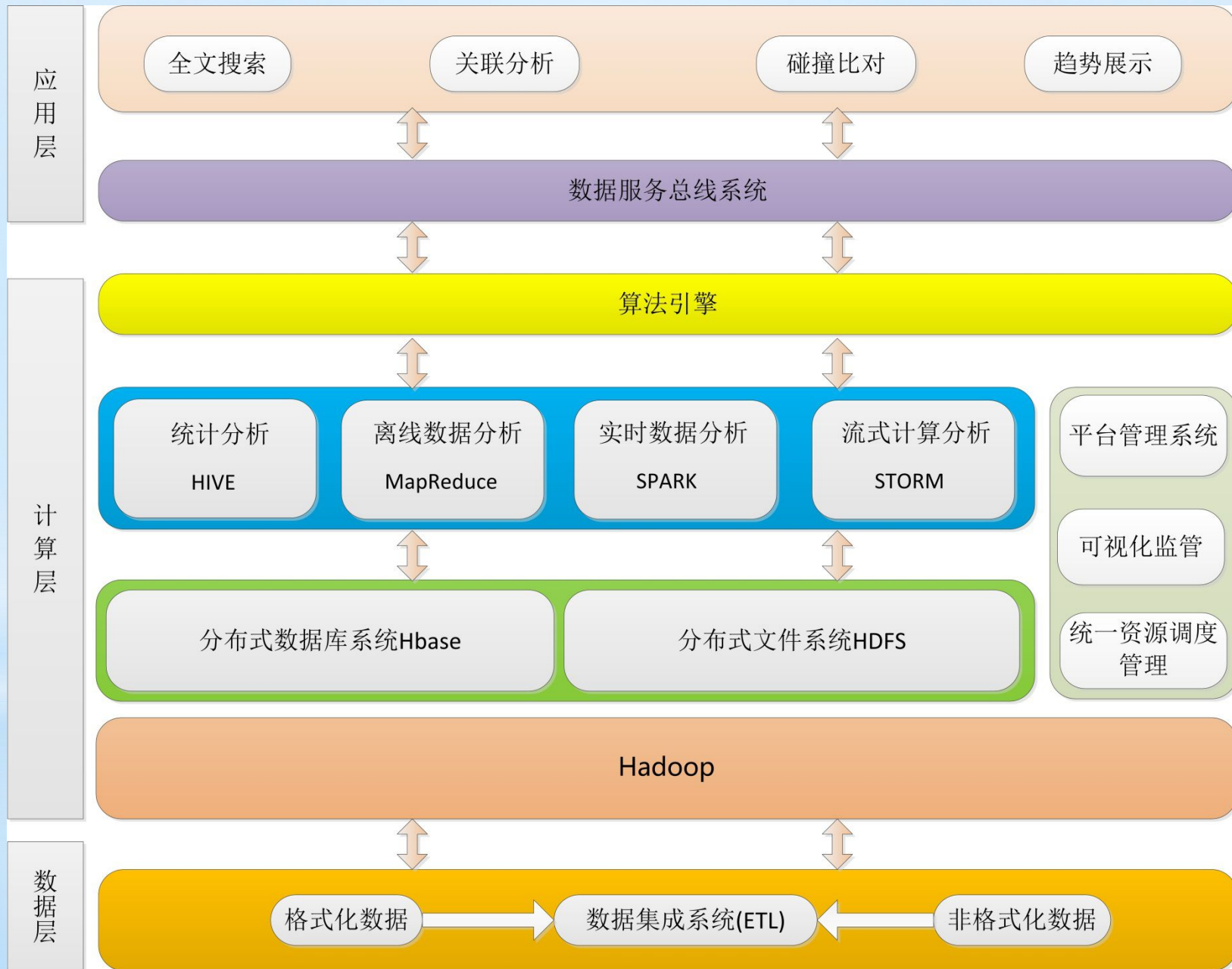
# 数据驱动开发模式创新

应用驱动开发以功能需求为导向

数据驱动开发以数据关系和应用为导向

以**环境保护**为例，如果以业务应用需求为核心进行开发，需要梳理几十种业务需求，建立几十种应用场景，开发几十种应用系统；如果以数据为核心，梳理业务应用和场景可以大大简化，并且可以发现新的业务模式。

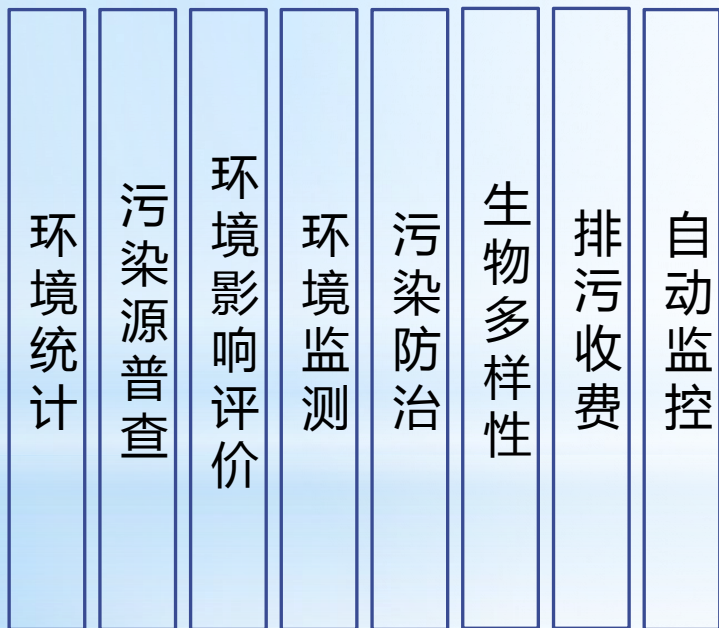
# 数据驱动环境信息资源中心架构设计



# 数据驱动环境管理转型

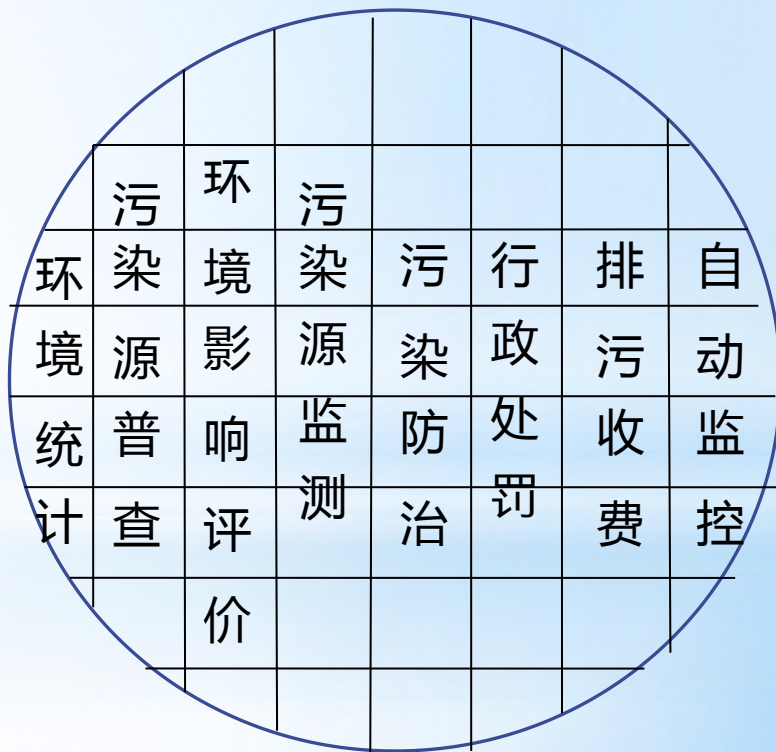
目前，人们所讨论的环保数据几乎都是“**条数据**”，即在某个环境管理业务呈链条状串起来的数据

条数据



未来，大数据发展的趋势是“**块**”上的“**条**”融合，即一个物理空间或者行政区域内形成的涉及环境管理对象各类数据的总和

块数据



“**条**”与“**块**”结合，产生数据协同效应，推动转型升级

# 数据驱动环境管理协同与合作

## 部门业务协同：

- ▶ 环保部门开放数据，实现数据交换与信息共享
- ▶ 解决环保部门业务不协同带来的矛盾与冲突
- ▶ 推进环保部门业务协同，提高管理与决策水平

## 统筹多部门合作（大部制改革）：

- ▶ 例如地下水监测与管理，**国土**建设10103个监测点，**水利**部建设10298个监测点，能够覆盖350万Km<sup>2</sup>
- ▶ 国土、水利与**环保**实现数据交换与信息共享，开展地下水合作监管模式
- ▶ 建立环境保护多部门联动合作机制

# 数据驱动环境监管模式创新

## 污染源自动监控：

- ▶ 通过对企业排污口安装监控设备，建立环保在线监控平台，实现对污染源的**实时监控**。
- ▶ 实时获取排污数据，监测污水排放是否超标，及时对企业进行监管。

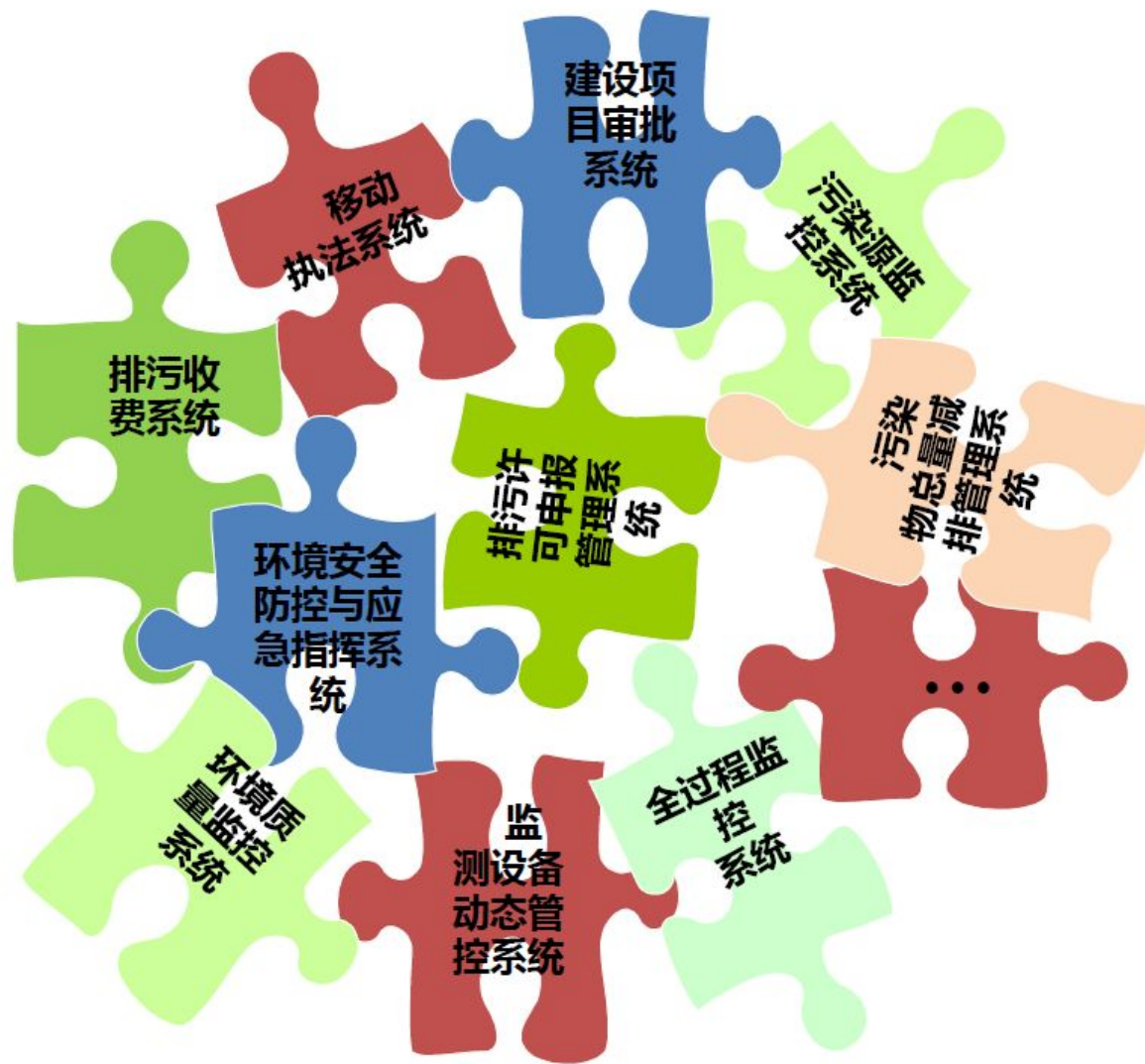
## 大数据监管模式：

- ▶ 多元数据采集相当于有了无数双眼睛，大数据分析增强预警预测能力。
- ▶ 通过投诉率、合规性、收益率、特征词等**大数据监管算法**，分析企业是否有偷拍漏排嫌疑，排污设施是否闲置。
- ▶ 对企业构建良好的**沟通反馈机制**，及时将监控数据反馈给污染企业，有助于构建长期、良性的数据质量保障工作，从而让企业形成自觉意识。
- ▶ 建立**企业环境信用评价体系**，通过数据分析计算企业的环境行为信息，按照企业环境信用评价指标，确定参评企业的环境信用等级，从而督促企业自觉履行环境保护法定义务和社会责任。
- ▶ 环评“未批先建”大数据监管平台。
- ▶ “水十条”、“气十条”大数据绩效评价体系。
- ▶ 大数据“环境监管与执法”评估体系。

# 数据驱动环境形势综合研判（一）

环境信息化系统多

系统间独立性很强



积累的数据量较大

数据间关联性很弱

# 数据驱动环境形势综合研判（二）

## 污染源全生命周期过程



## 需采集关联以下系统数据



数据多而分散，具有交叉性，全生命周期数据获取困难乃至无法全部获取



# 数据驱动环境管理体系创新

建立区域化大气污染管理体系

建立流域化水污染管理体系

建立典型区划土壤污染管理体系

建立功能区划生态保护管理体系

以数据开放和信息  
共享为基础，建立  
新型环境管理体系



环境保护需要

**大众关注、万众参与**

**信息公开、数据开放**

如果政府部门能够开放数据，将推动经济持续高速增长

**数据驱动经济效益**

**GDP>8%**

**谢谢**