



CHINA  
DATA  
ANALYST  
SUMMIT

CDA 数据分析师  
www.cda.cn

# 数据可视化-无限可能的艺术

ADD YOUR POWERPOINT TITLE HERE

演讲人：贺华 IBM资深软件技术顾问

跨界互联  
数聚未来

第四届中国数据分析师行业峰会  
CHINA DATA ANALYST SUMMIT

北京 中国大饭店 2017.07

# 内容提要

- **大数据时代面临的挑战**
- **IBM业务分析解决方案**
- **可视化案例分享**



# THE BIG BANG THEORY

TIME BEGINS

ONE SECOND

PRESENT DAY

Time  $10^{-43}$  sec.  
Temperature

$10^{-32}$  sec.  
 $10^{27}$ °C

$10^{-6}$  sec.  
 $10^{13}$ °C

3 min.  
 $10^8$ °C

300,000 yrs.  
 $10,000$ °C

1 billion yrs.  
 $-200$ °C

15 billion yrs.  
 $-270$ °C

**1** The cosmos goes through a superfast "inflation," expanding from the size of an atom to that of a grapefruit in a tiny fraction of a second

**2** Post-inflation, the universe is a seething, hot soup of electrons, quarks and other particles

**3** A rapidly cooling cosmos permits quarks to clump into protons and neutrons

**4** Still too hot to form into atoms, charged electrons and protons prevent light from shining; the universe is a superhot fog

**5** Electrons combine with protons and neutrons to form atoms, mostly hydrogen and helium. Light can finally shine

**6** Gravity makes hydrogen and helium gas coalesce to form the giant clouds that will become galaxies; smaller clumps of gas collapse to form the first stars

**7** As galaxies cluster together under gravity, the first stars die and spew heavy elements into space; these will eventually form into new stars and planets

Quarks

Neutron

Hydrogen nucleus

Hydrogen atom

Protogalaxy

Electron

Proton

Helium nucleus

Helium atom

Galaxy



# 大数据的特征 4个V



# BIG DATA



**VOLUME**  
DATA SIZE

量级



**VELOCITY**  
SPEED OF CHANGE

时效



**VARIETY**  
DIFFERENT FORMS  
OF DATA SOURCES

复杂



**VERACITY**  
UNCERTAINTY OF  
DATA

精准

# 如何理解大数据？

- Big Data = 大数据 + 大数据的处理方案 (**存储+计算+展现**)
- “Big Data” 泛指数据集的大小超过了通常的数据库，数据管理软件所能获取，存储，管理和分析范围的结构化与非结构化数据
- “Big Data” 是一个相对的概念，不能从绝对意义上指明某一个数值来定义，而且这个相对的标准根据所处的行业和客户的不同千差万别。



## Big Data解决方案分层

应用--访问及展现

计算框架

数据存储



# 大数据和分析功能在企业的主要应用模式



## 可供采取行动的全方位客户洞察

如何才能最深入地了解客户和潜在客户，以期让每一次交互更有意义，从而把客户转化为终生的品牌拥趸？



## 商品优化

是否能够预测应该购入哪种商品以及如何在各个地点上架和定价，从而优化我的业务？能够在客户服务与利润最大化之间实现平衡吗？



## 数据和模式探索

怎样才能快速找到、可视化及理解所有数据，以改进决策和客户服务？



## 基于位置的服务

如何分析客户的购物模式，以提供有针对性的信息及更好的店内服务，并尽早预防欺诈？



## 数据仓库现代化

如何降低现有IT基础架构的成本，同时还要为各个部门提供更高程度的信息利用率？借助于Hadoop、分析工具以及分栏（columnar）引擎，是否能够找到更好的性价比选择？

# 第5个V: Value 价值

Value

noun

In the research  
association

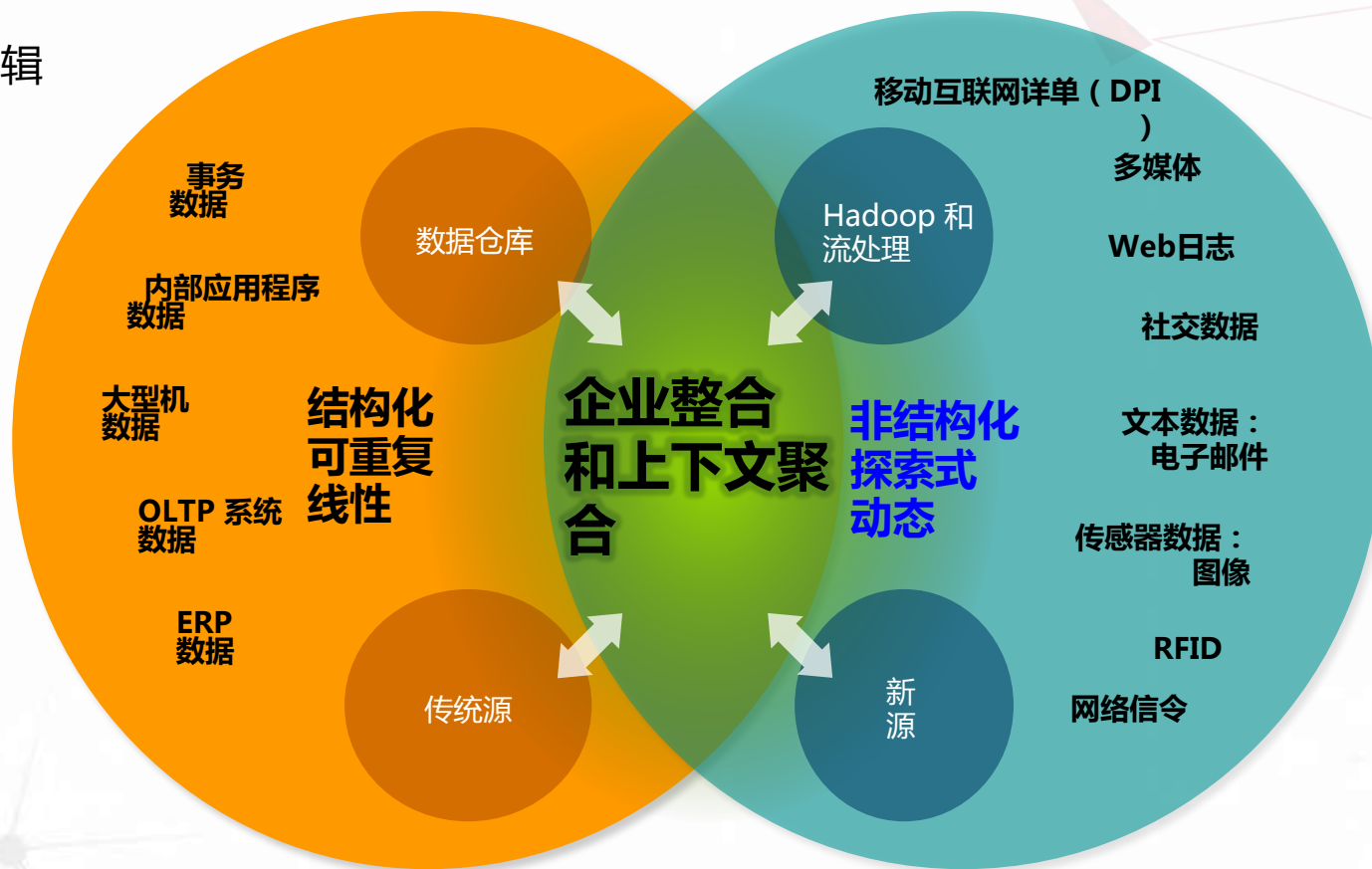




# 大数据如何释放新洞察、如何创建新机遇？

## 传统方法

结构化、分析、逻辑

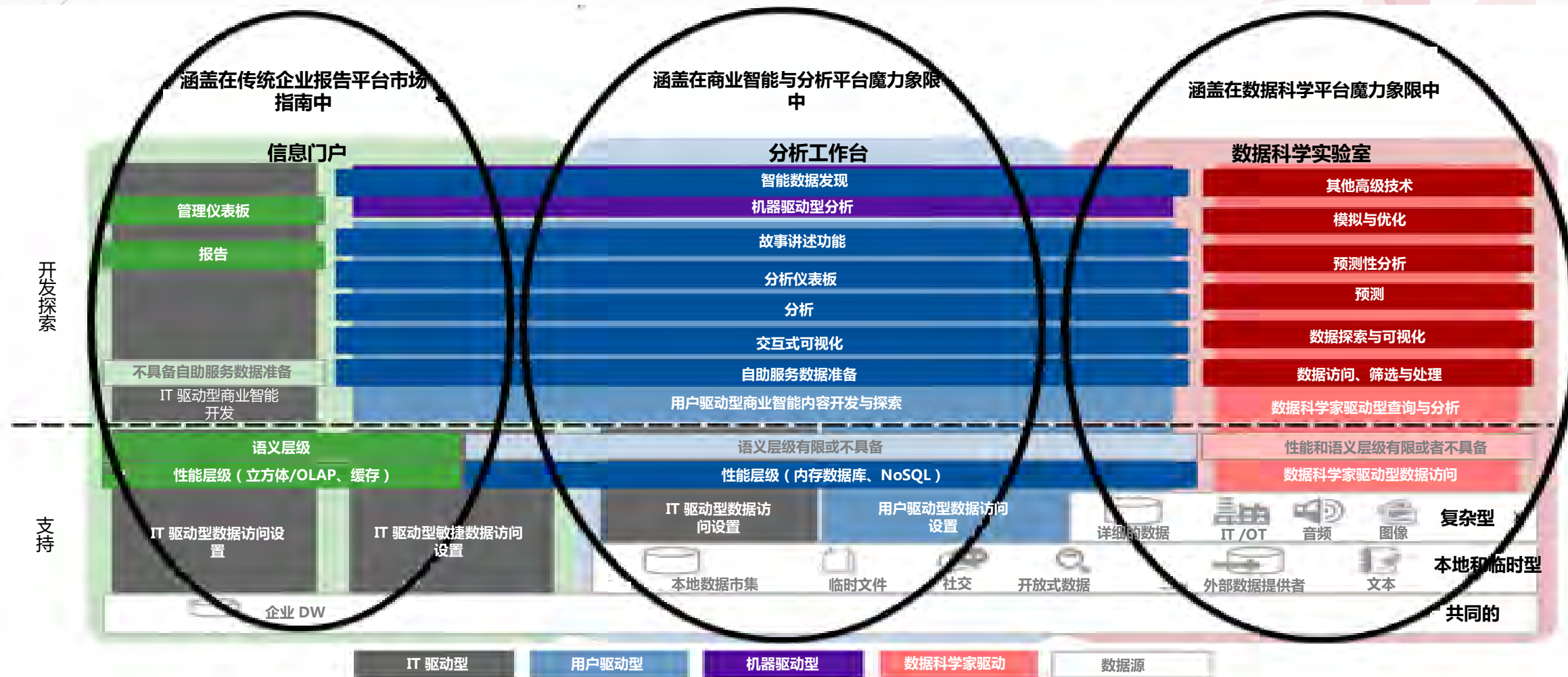


## 新方法

创新、全面考虑、直观



# Gartner 现代商业智能和分析参考体系结构



来源：Gartner “如何实施现代商业智能和分析平台” - <https://www.gartner.com/technology/media-products/newsletters/ibm-cognos/1-3K3NKKK/index.html>

# 欢迎进入认知时代 这是一个新的技术时代、新的商务时代



越来越多的软件公司会把分析能力嵌入到他们的业务应用程序之中。然而讽刺的是，很多客户仍然没有意识到这样做带来的影响和附加的业务价值。

Source: Hurwitz & Associates, Predicting 2016 What's Old is New Again

跨界互联  
数聚未来

第四届中国数据分析师行业峰会  
CHINA DATA ANALYST SUMMIT



CIDA 数据分析师  
www.cda.cn



# 企业级分析如何满足业务用户需求

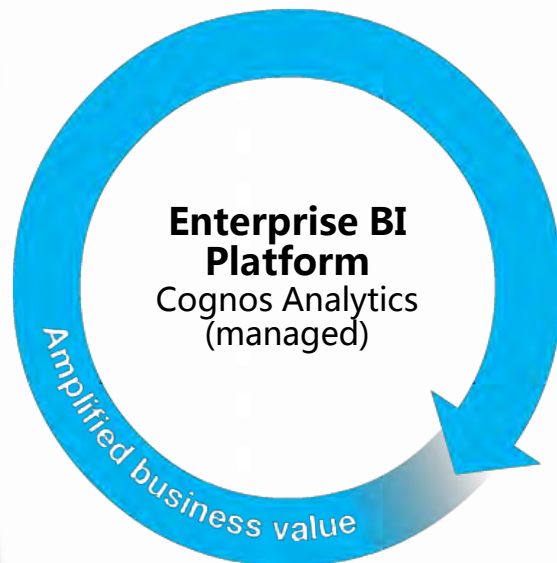
## 业务专家

定制分析内容，创建仪表盘、即席查询报表、数据模块



## 业务领导

查看静态及交互式报表，向其它下属提问



业务人员，业务分析师  
创建与分享报表及仪表盘

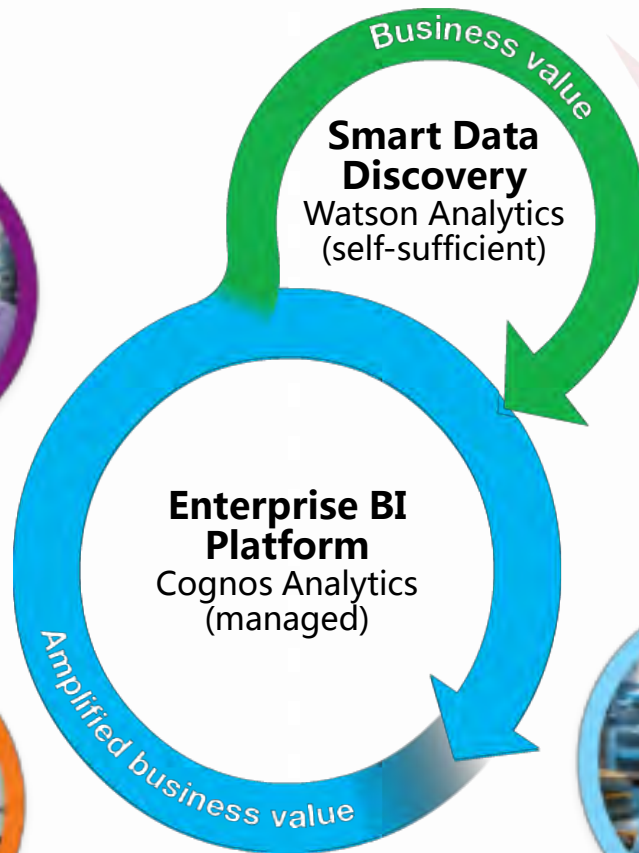


# 智能数据探索拓展对企业数据的理解

**业务专家**  
通过自然语言探索数据，发现预测性因素，并通过仪表盘和图表可视化探索



**业务领导**



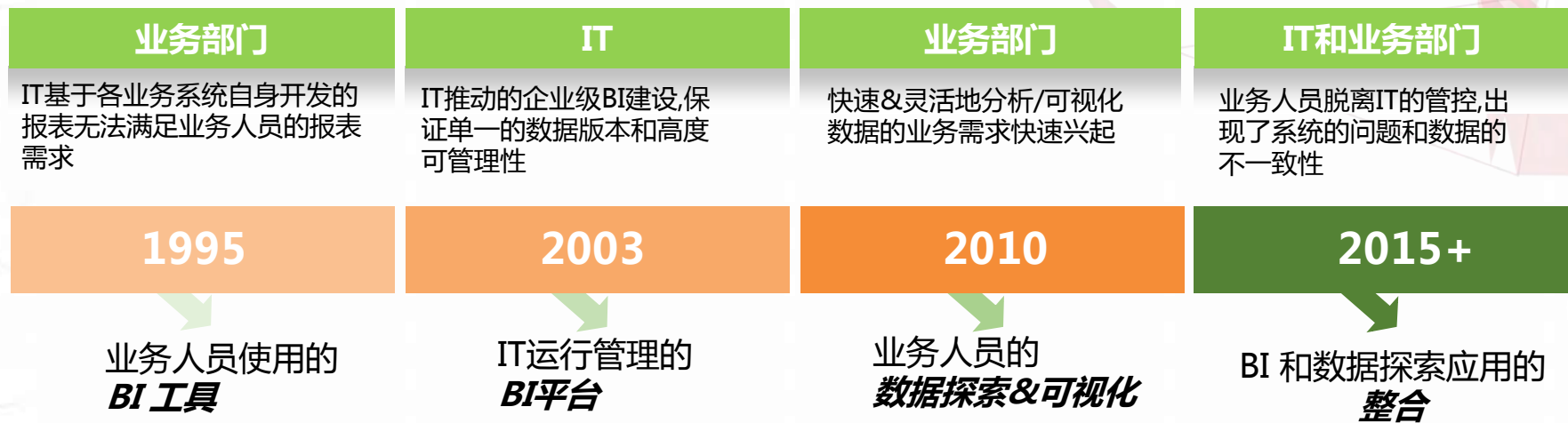
相关见解返回到企业分析平台

**业务人员，业务分析师**





# 商业智能被业务部门和IT交替推动着向前发展



**Gartner prediction:** Through 2016, less than  
of self-service business intelligence initiatives will be governed  
sufficiently to prevent inconsistencies that adversely affect the  
business.

10%

Source: Gartner "Predicts 2015: Power Shift in Business Intelligence and Analytics Will Fuel Disruption", Josh Parenteau, Neil Chandler, Rita L. Sallam, Douglas Laney, Alan D. Duncan, 21 November 2014

# Cognos Analytics将敏捷BI融合到企业BI中,使业务用户自主分析和企业IT管控间保持平衡

## 寻求完美平衡

在业务部门级的自由分析和在企业级的安全、管控和监督的需求之间





# 自助服务分析是一个系列



业务用户

90% 普通用户  
(高管、经理、一线工作人员)

10% Power 用户  
负责分析数据



60%  
数据消费者



30%  
数据开发者



8%  
数据分析师



2%  
数据科学家

“自助服务分析的弊端，就是它需要实现标准化。”

“许多已经部署自助服务分析的公司已经被海啸般大量的矛盾冲突报告、报表集市 (spreadmart)、脱节的报告系统以及其他数据孤岛所淹没。”

Source: Chart data and quotes from “Eckerson Group: A Reference Architecture for Self-Service Analytics” , Wayne Eckerson, Barry Devlin, September 2016

# Cognos Analytics 支持一系列自助服务功能



90% 普通用户  
(高管、经理、一线工作人员)

10% Power 用户  
负责分析数据



60%  
数据消费者



30%  
数据开发者



8%  
数据分析师



2%  
数据科学家

静态  
报告

交互式  
报告

个性化  
报告

仪表板和  
案例

数据模块和临时报告

数据模式和专业报告

- 业务用户从单纯使用数据发展至扩展和创造数据 (成为数据开发者)

Source: Chart data from "Eckerson Group: A Reference Architecture for Self-Service Analytics", Wayne Eckerson, Barry Devlin, September 2016

跨界互联  
数聚未来

第四届中国数据分析师行业峰会  
CHINA DATA ANALYST SUMMIT



CIDA 数据分析师  
www.cda.cn



# Cognos Analytics – 针对统一环境的框架

## 自助服务 (仪表盘, 案例)

- 数据**开发** (临时)
- 主要用例
  - 回答问题
  - 头脑风暴/构想与概念的社会化
  - 构建原型



业务线用户

## 管理报告

- 数据呈现
- 通过管理下列内容实现增值：
- 安全和治理
  - 输出完整性
  - 输出的有效分配



熟练的用户 (Power/IT)

## 平台

- 支持安全性、可扩展性和完整性的企业级架构
- 为自助服务和管理报告接口提供一致性

# Cognos Analytics – 针对统一环境的框架

## 自助服务功能 (仪表盘, 案例)

- 上传数据
- 轻量级数据建模
- 创建仪表盘
- 搜索
- 直观导航
- 互动式输出



业务线用户

## 管理报告功能

- 调度
- 爆发
- 事件监控
- 离线报告
- 移动报告
- SDK
- 混合服务
- 可扩展的可视化 (RAVE)
- 参数化报表
- 模版
- 等等...



熟练的用户 (Power/IT)

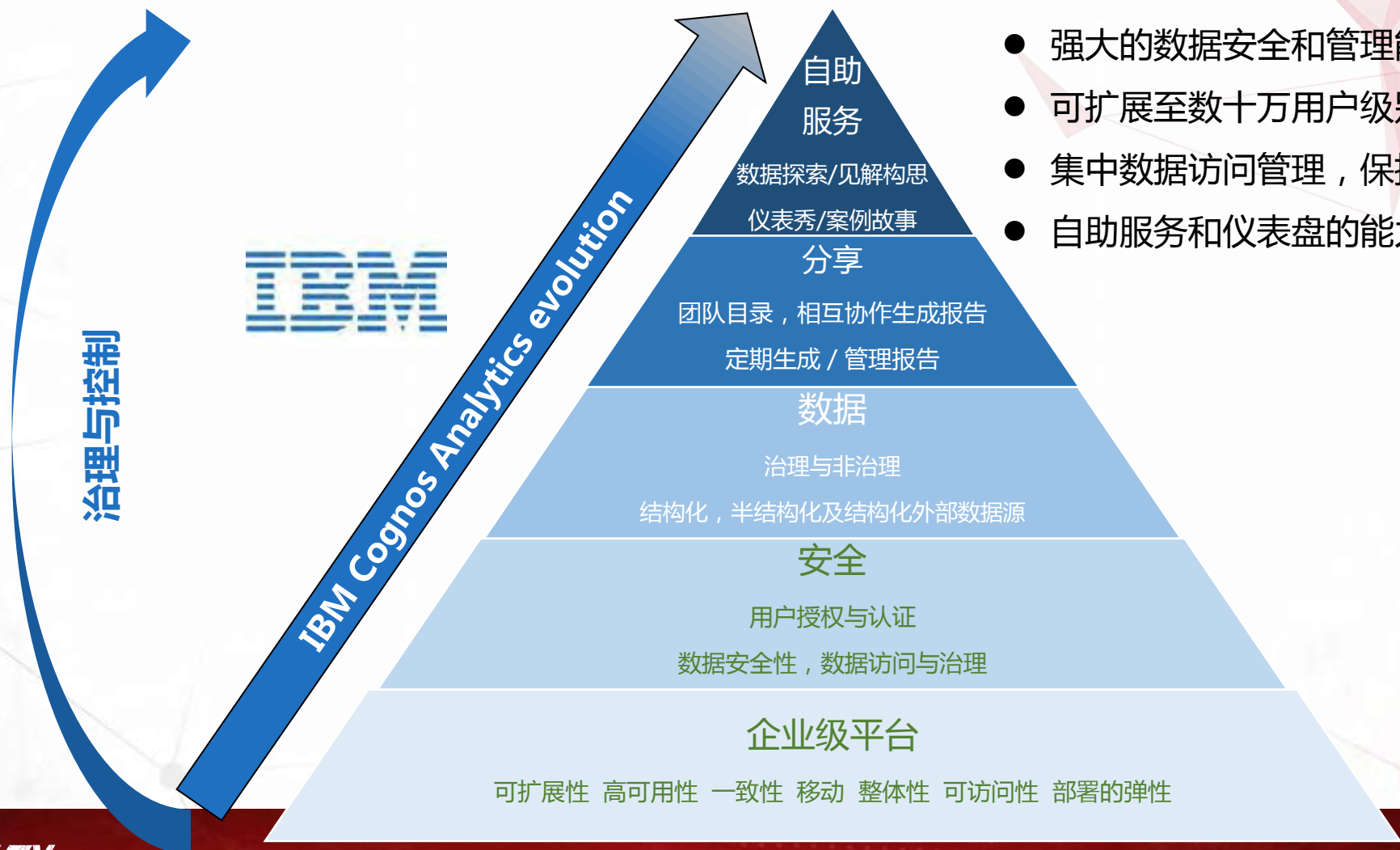
## 元数据层 平台

- 面向数据源的开放式访问
- 一般业务模式

- 具有成本效益的可扩展性
- 以服务为导向的体系结构 - 建立在开放标准之上



# Cognos Analytics提供企业级BI与分析能力



- 强大的数据安全和管理能力
- 可扩展至数十万用户级别的能力
- 集中数据访问管理, 保持分析一致性
- 自助服务和仪表盘的能力

# 为了更好的在自助分析和IT管控之间获得平衡， IBM对商业智能做了重新定义

## Cognos Analytics完全重新定义的BI用户体验

Dramatically increase productivity for departmental and enterprise reporting

加速  
**业务建模**  
及性能

重新定义  
**制作报表**  
和分析

重新定义  
**BI应用**  
在各种设备上的体验

+ **IBM Watson Analytics** 智能数据探索

跨界互联  
数聚未来

第四届中国数据分析师行业峰会  
CHINA DATA ANALYST SUMMIT



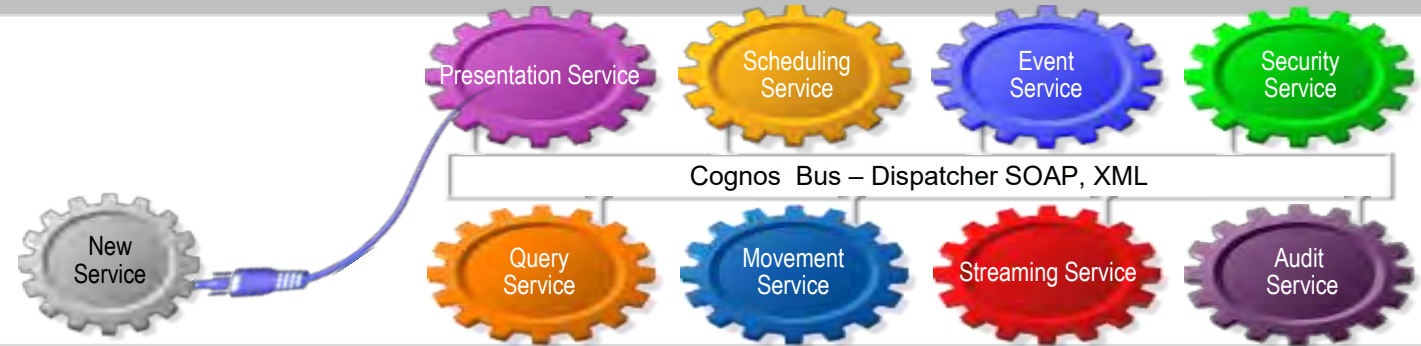
CIDA 数据分析师  
www.cda.cn



# Cognos Analytics同时提供企业级BI+敏捷BI能力



- 统一的元数据
- 统一的安全性
- 统一优化的查询服务
- 单一的系统管理
- 负载容错
- 开放的API(数据和管理)



## 公共的企业BI模型



# IBM可视化分析引擎RAVE

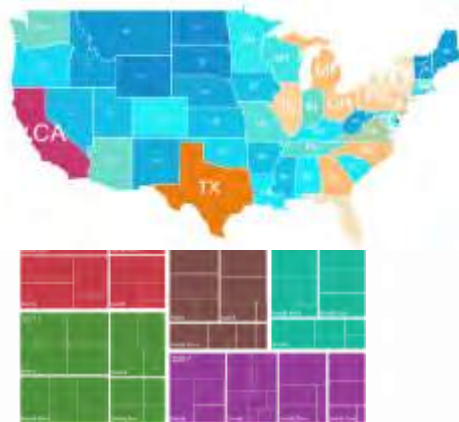
## A paradigm shift for delivering value to users

### 挑战:

- 新的可视化需求随着大数据的出现在不断增加
- 每个需求都是特别的,难以直接满足

### 解决方案:

- 提供大量的可视化种类,也就是意味着可以方便地定制.

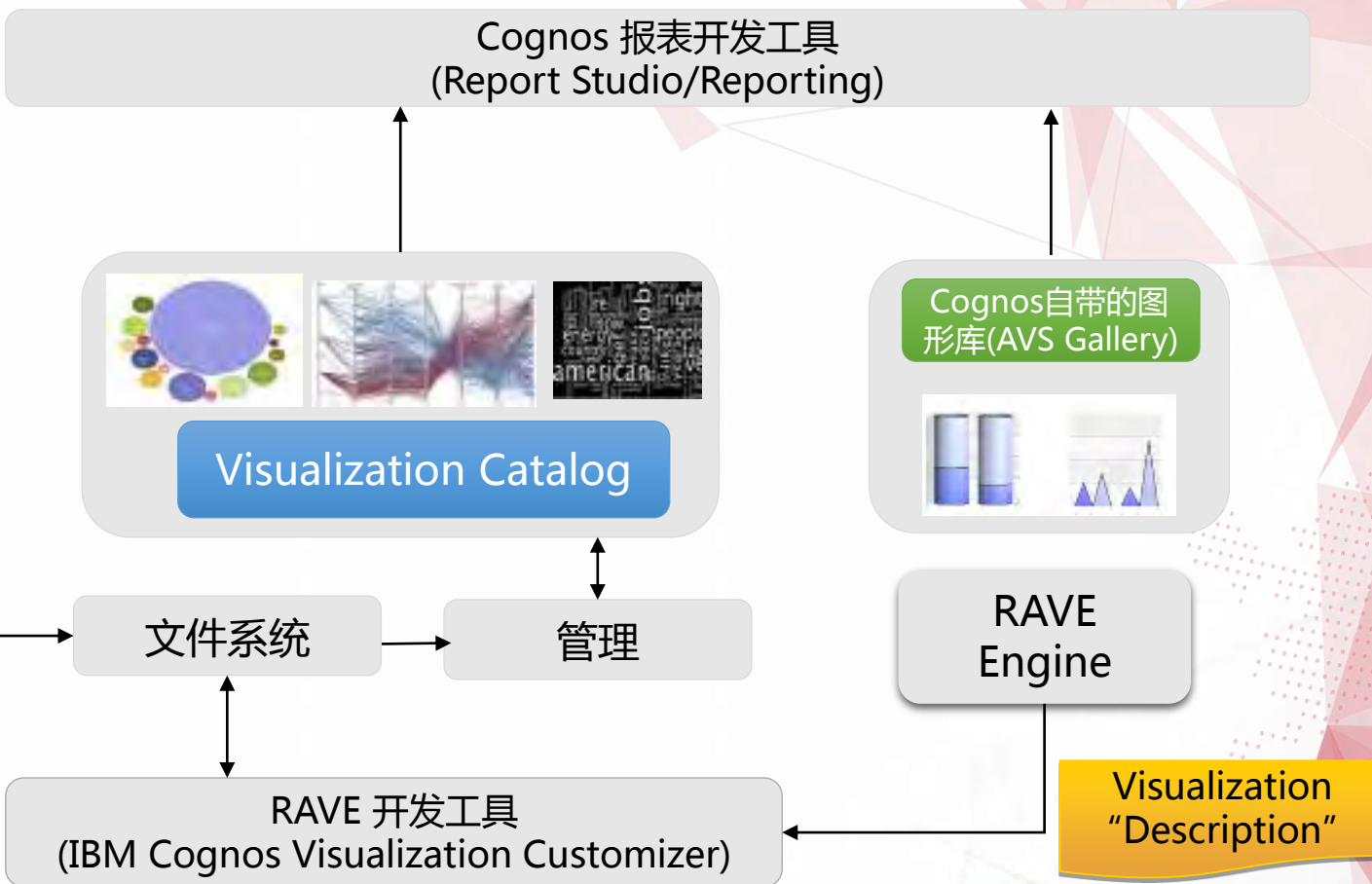
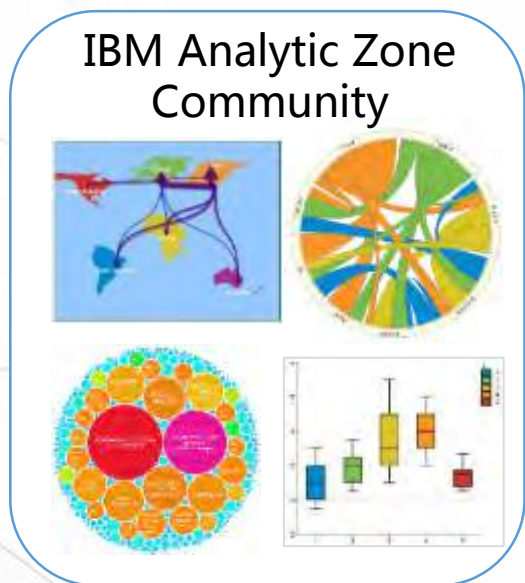




# IBM RAVE图形引擎提供了大量的可视化示例

The screenshot displays the IBM RAVE website interface. At the top left is the RAVE logo, and at the top right is a blue button labeled "Get Rave Now". Below the logo is a navigation menu with links for Home, Visualizations, Media, Downloads, Blog, Apps, Learn, and Workbench. The main heading is "Visualizations" with a subtext "Using RAVE 1.8.0.10 (Schema 7.2)". A search bar is present with the text "Search Visualizations". On the left side, there is a "Filters" section with a "Clear All" button. Under "Provider", there are checkboxes for Rave (58), Research (9), PureScale (3), BI (18), FPM (2), Hackday (9), Experimental (1077), and Rave Development (29). Under "Chart Type", there are checkboxes for area (63), edge (151), interval (550), line (141), path (9), point (521), polygon (29), schema (4), schema:boxplot (20), and text (36). The main content area features a 3x3 grid of nine different visualization examples, each with a red "Rave" banner in the top right corner. The visualizations include: a combined bar and line chart; a circular bubble chart; a line chart with a shaded area; a line chart with multiple data series; a network graph; a horizontal bar chart; a horizontal bar chart with error bars; a grouped bar chart; and a boxplot.

# RAVE图形引擎与Cognos Analytics



<https://www.ibm.com/blogs/analytics-zone/>

跨界互联  
数聚未来

第四届中国数据分析师行业峰会  
CHINA DATA ANALYST SUMMIT



CIDA 数据分析师  
www.cda.cn



# RAVE 开发工具

## - IBM Cognos Visualization Customizer

The screenshot displays the IBM Cognos Visualization Customizer application. The interface is divided into several sections:

- Left Panel:** A code editor showing JSON configuration for a visualization. The visible code includes:

```
428     "family": "'Arial Unicode  
429     "size": "10pt"  
430   }  
431 },  
432   "titleStyle":  
433   {  
434     "fill": "#808080",  
435     "align": "middle",  
436     "font":  
437     {  
438       "family": "'Arial Unicode  
439       "size": "10pt"  
440     },  
441     "padding":  
442     {  
443       "bottom": 5  
444     }  
445   },  
446   "legendEntryStyle":  
447   {  
448     "padding": 0  
449   },  
450   "method": "swatch"  
451 }  
452 ],  
453 "resource":  
454 [  
455   {  
456     "id": "id1",  
457     "type": "geo/json/text",  
458     "URIs":  
459     [  
460       "embedded:///resource/*[\\"id  
461     ],  
462     "embeddedData": "{\\"type\\":\\"Fea  
463
```
- Top Panel:** Contains tabs for "RAVE", "Chart Type Definition", and "Thumbnails". Below the tabs are controls for "Draw Chart" (checked), "Validate" (checked), "Animation" (set to "None"), and "Background" (set to "White"). There are also dropdown menus for "Type" (set to "element"), "SubType" (set to "(all)"), and "Info".
- Right Panel:** A large preview window showing a world map with three pie charts overlaid on different regions (North America, Europe, and Asia).