DTCC

2017第八届中国数据库技术大会

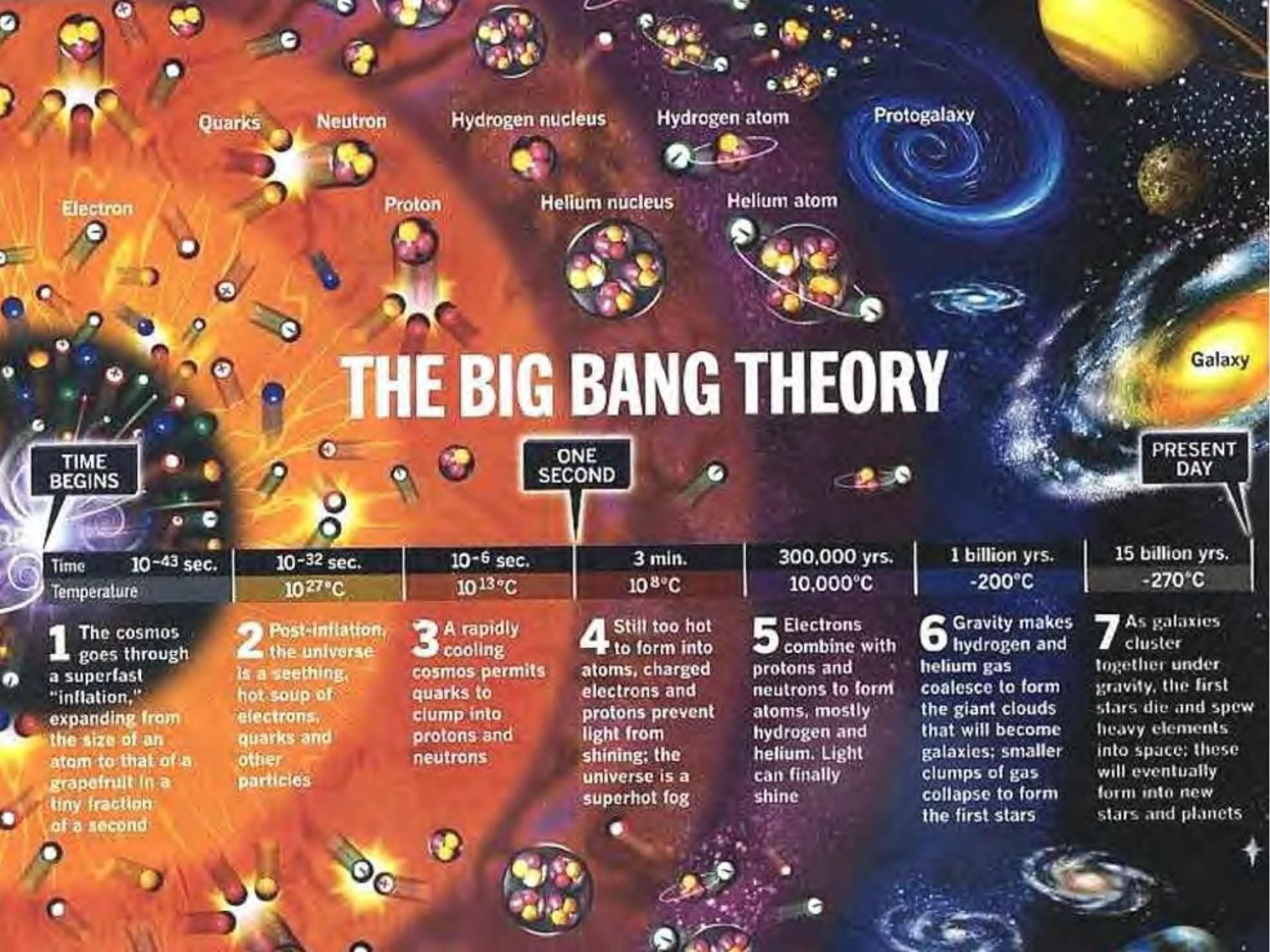
DATABASE TECHNOLOGY CONFERENCE CHINA 2017

数据可视化-无限可能的艺术

贺华 IBM资深软件技术顾问

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.	10^{-32} sec.	10^{-6} sec.	3 min.	300,000 yrs.	1 billion yrs.	15 billion yrs.
Temperature		10^{27}°C	10^{13}°C	10^{8}°C	10,000°C	-200°C	-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

大数据的特征 4个V


BIG DATA
VOLUME
DATA SIZE
VELOCITY
SPEED OF CHANGE
VARIETY
DIFFERENT FORMS OF DATA SOURCES
VERACITY
UNCERTAINTY OF DATA
量级
时效
复杂
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

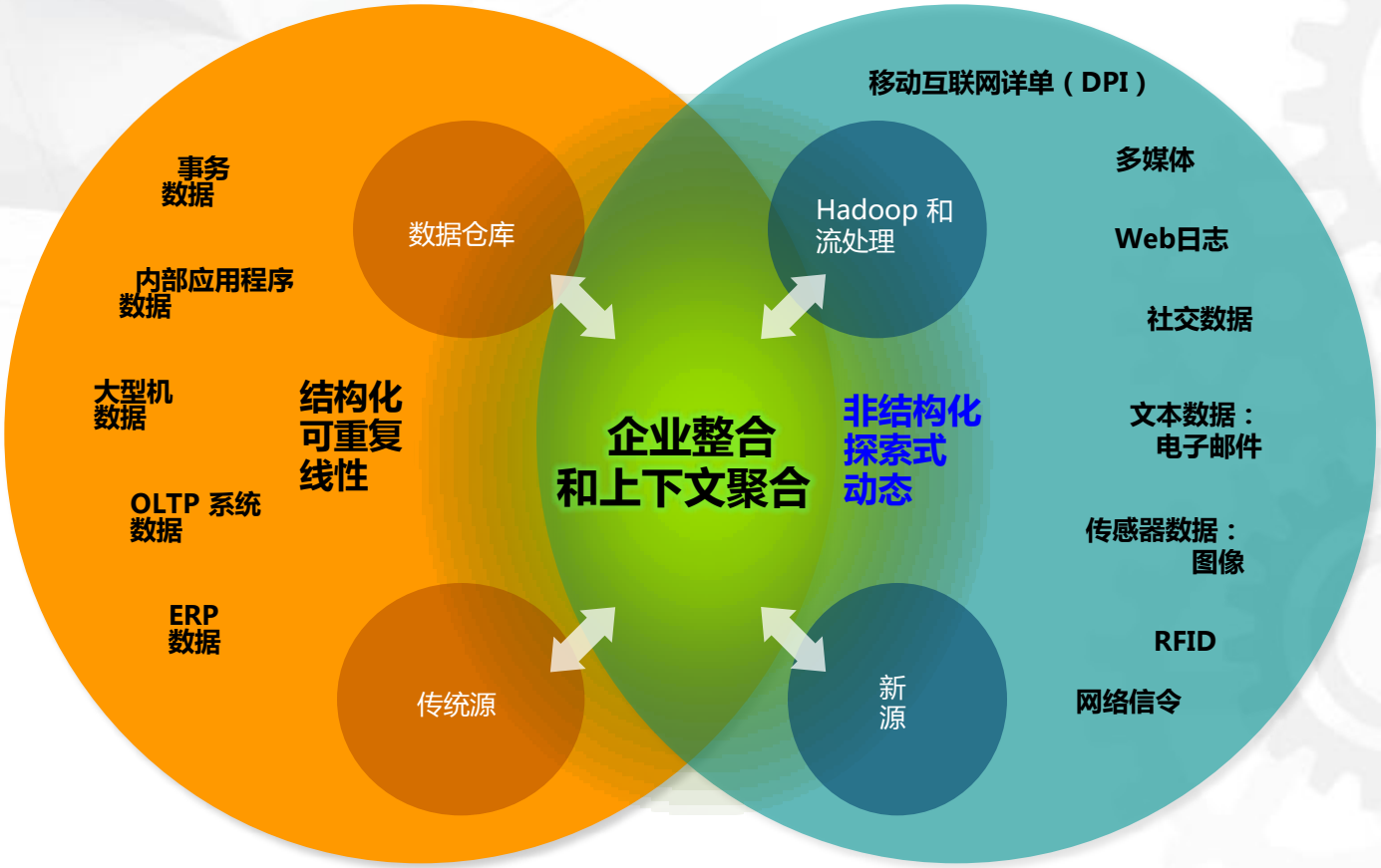
如何降低现有IT基础架构的成本，同时还要为各个部门提供更高程度的信息利用率？借助于Hadoop、分析工具以及分栏（columnar）引擎，是否能够找到更好的性价比选择？

第5个V: Value 价值

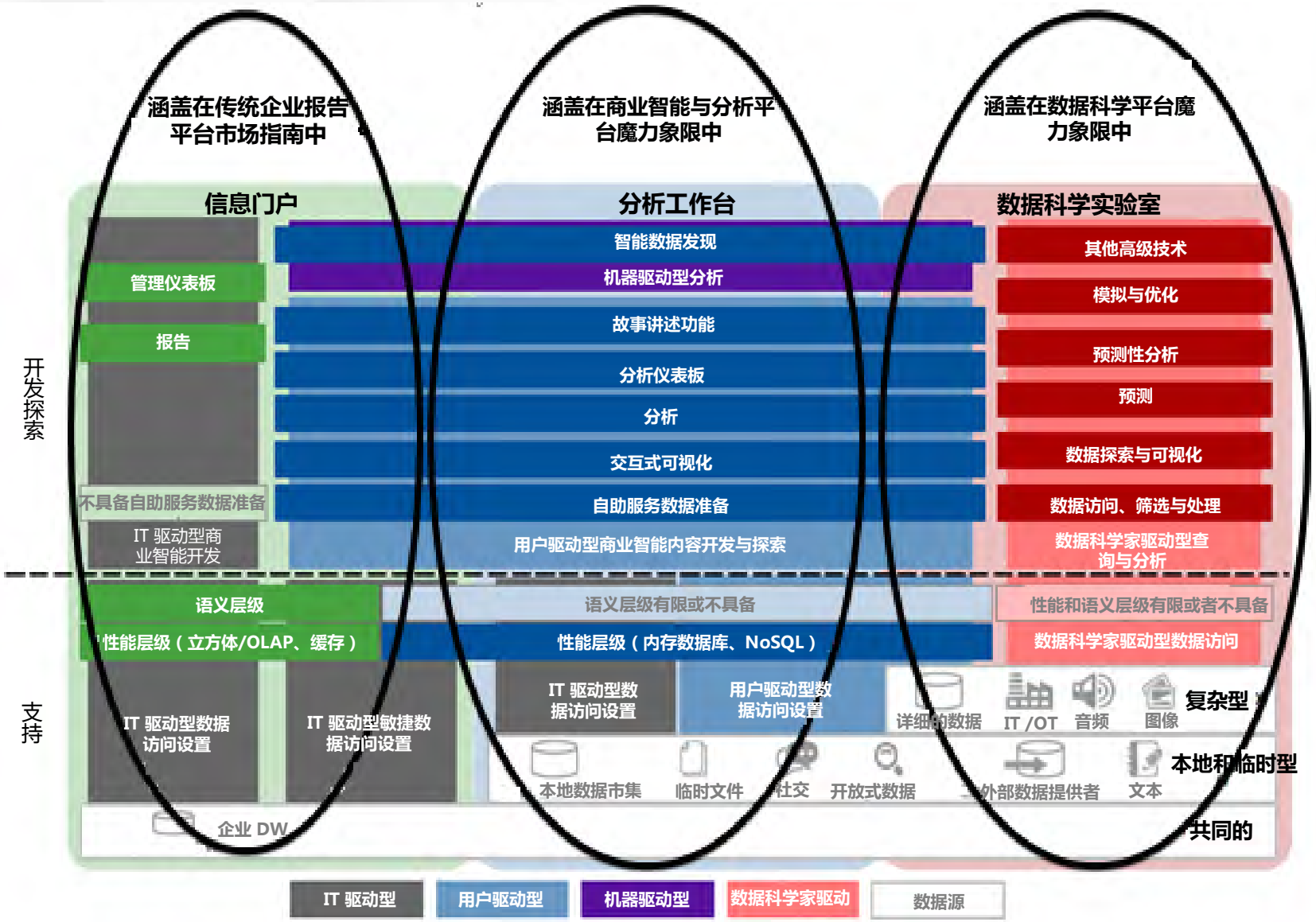
大数据如何释放新洞察、如何创建新机遇？

传统方法
结构化、分析、逻辑

新方法
创新、全面考虑、直观

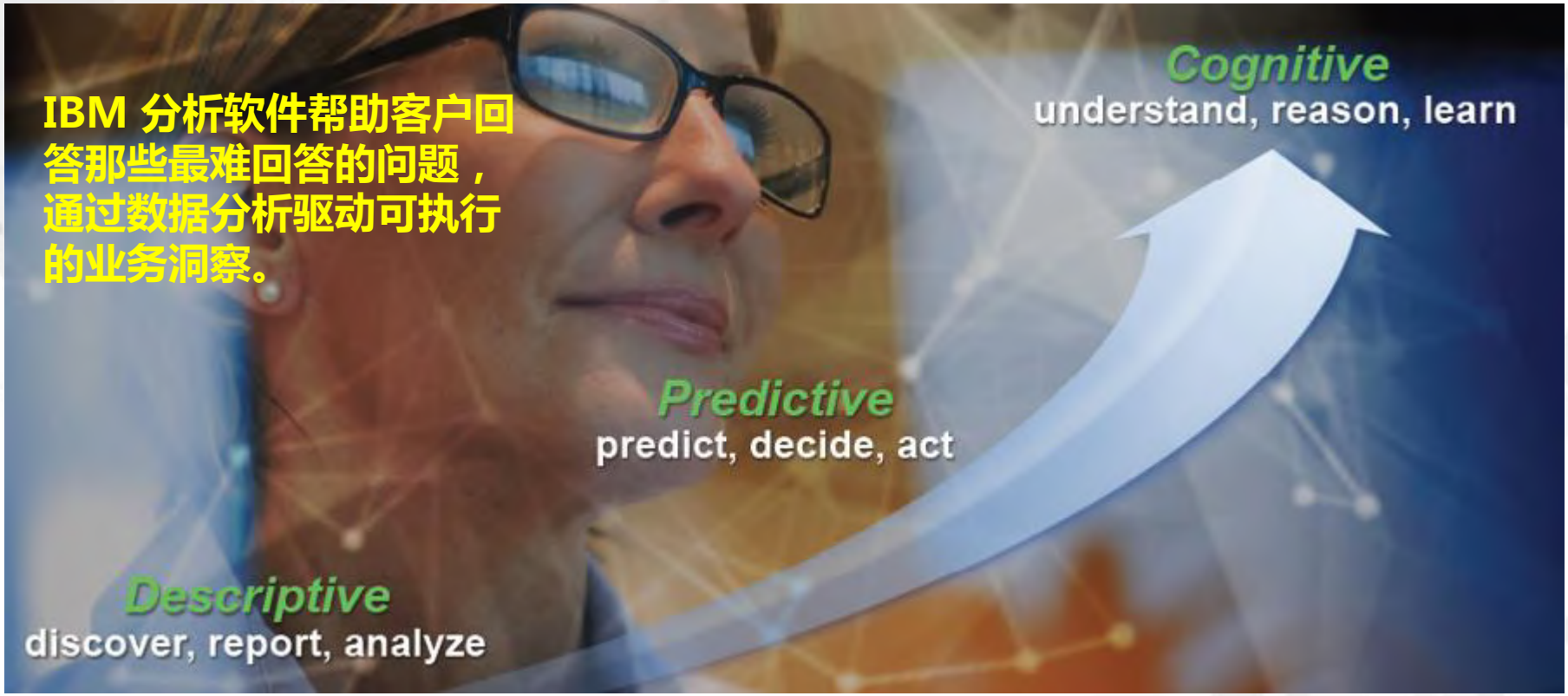


Gartner 现代商业智能和分析参考体系结构



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/1-3K3NKKN/index.html

Welcome to the Cognitive Era.
A new era in technology, a new era in business.



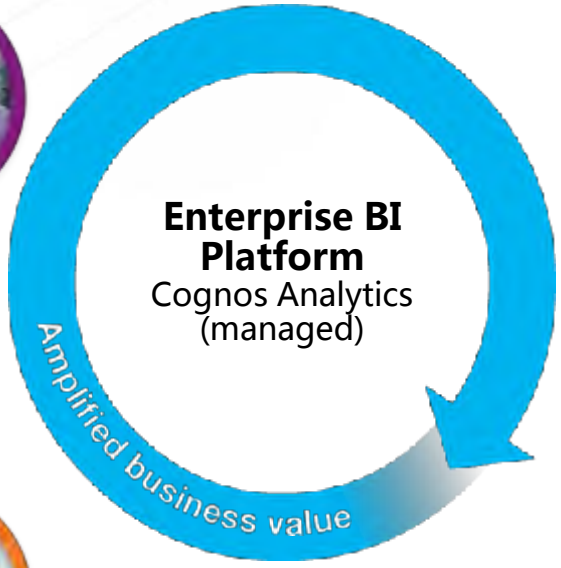
越来越多的软件公司会把分析能力嵌入到他们的业务应用程序之中。然而讽刺的是，很多客户仍然没有意识到这样做带来的影响和附加的业务价值。

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

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

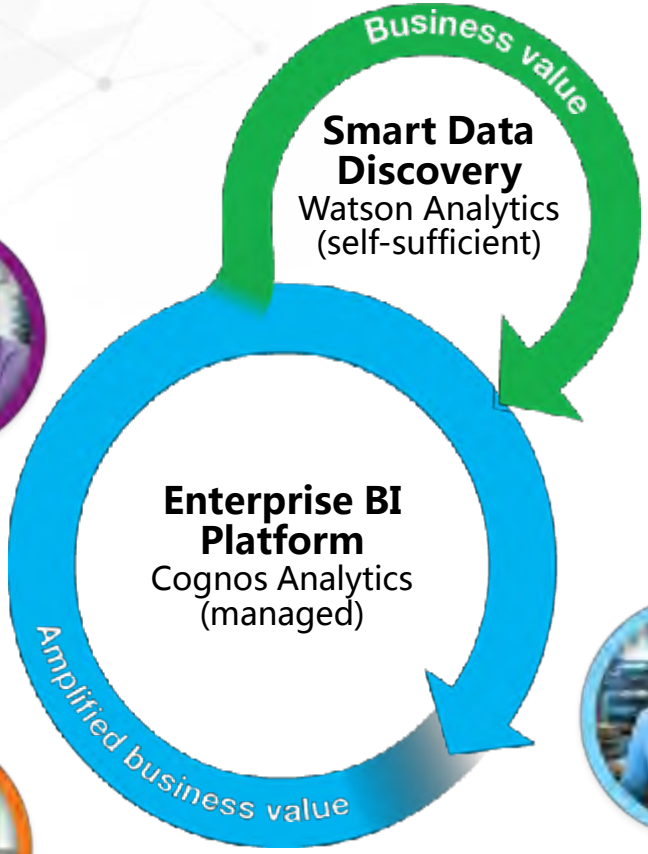


业务人员，业务分析师
创建与分享报表及仪表盘

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

1995	2003	2010	2015+
业务部门	**IT**	**业务部门**	**IT和业务部门**
IT基于各业务系统自身开发的报表无法满足业务人员的报表需求	IT推动的企业级BI建设,保证单一的数据版本和高度可管理性	快速&灵活地分析/可视化数据的业务需求快速兴起	业务人员脱离IT的管控,出现了系统的问题和数据的不一致性
业务人员使用的 ***BI 工具***	IT运行管理的 ***BI平台***	业务人员的 ***数据探索&可视化***	BI 和数据探索应用的 ***整合***

Gartner prediction: Through 2016, less than **10%** of self-service business intelligence initiatives will be governed sufficiently to prevent inconsistencies that adversely affect the business.

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

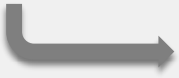
自助服务分析是一个系列

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

业务用户

静态报告 *互动式报告* *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

Cognos Analytics – 针对统一环境的框架

自助服务
（仪表板，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)

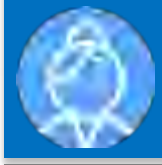
平台

- 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
（仪表板，案例）

- 上传数据
- 轻量级数据建模
- 创建仪表板
- 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...

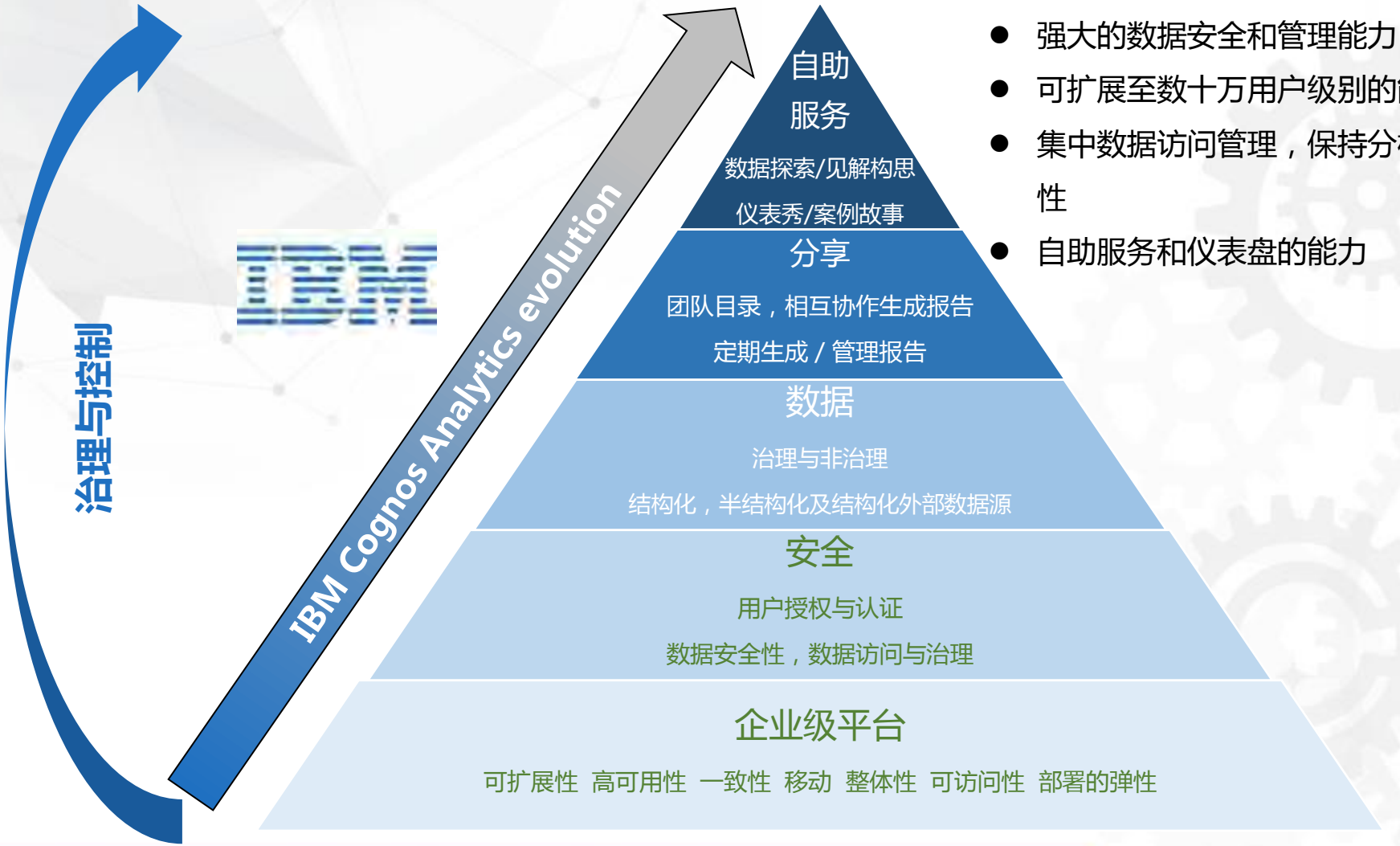
熟练的用户 (Power/IT)

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

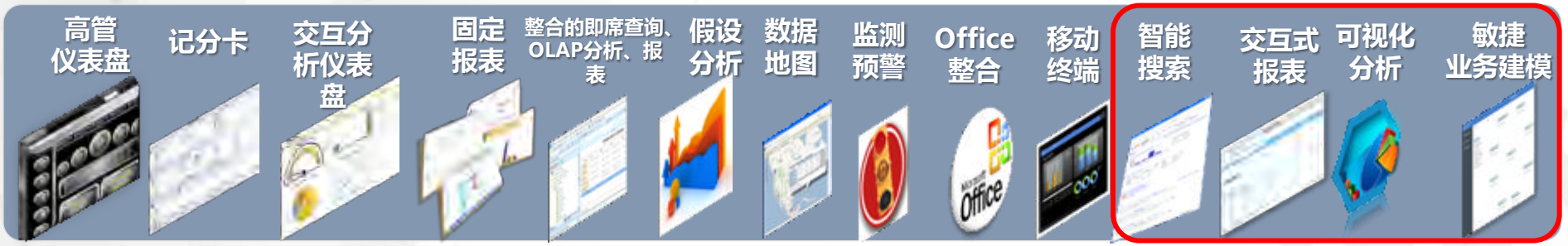
Cognos Analytics完全重新定义的BI用户体验

Dramatically increase productivity for departmental and enterprise reporting

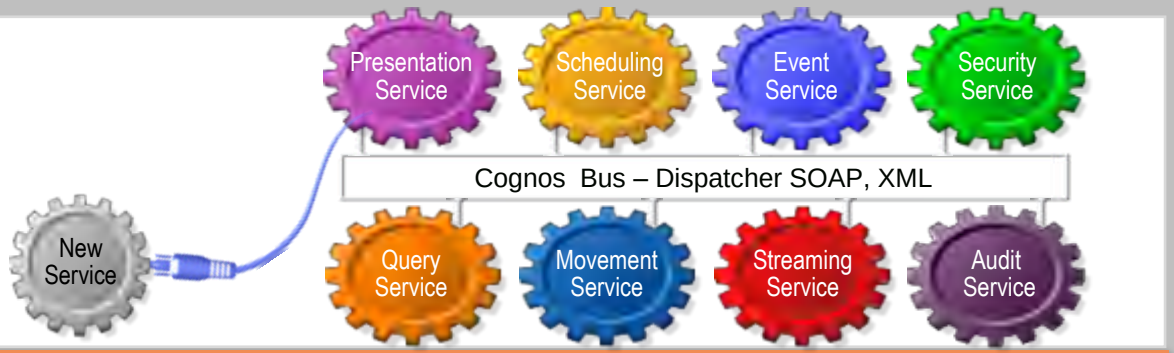
加速	重新定义	重新定义
业务建模	**制作报表**	**BI应用**
及性能	和分析	在各种设备上的体验

\+ IBM Watson Analytics 智能数据探索

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

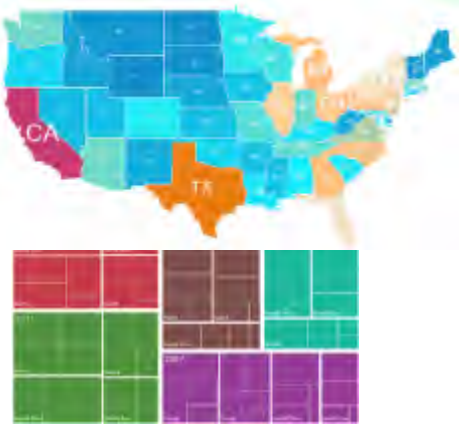
IBM可视化分析引擎RAVE
A paradigm shift for delivering value to users

挑战:

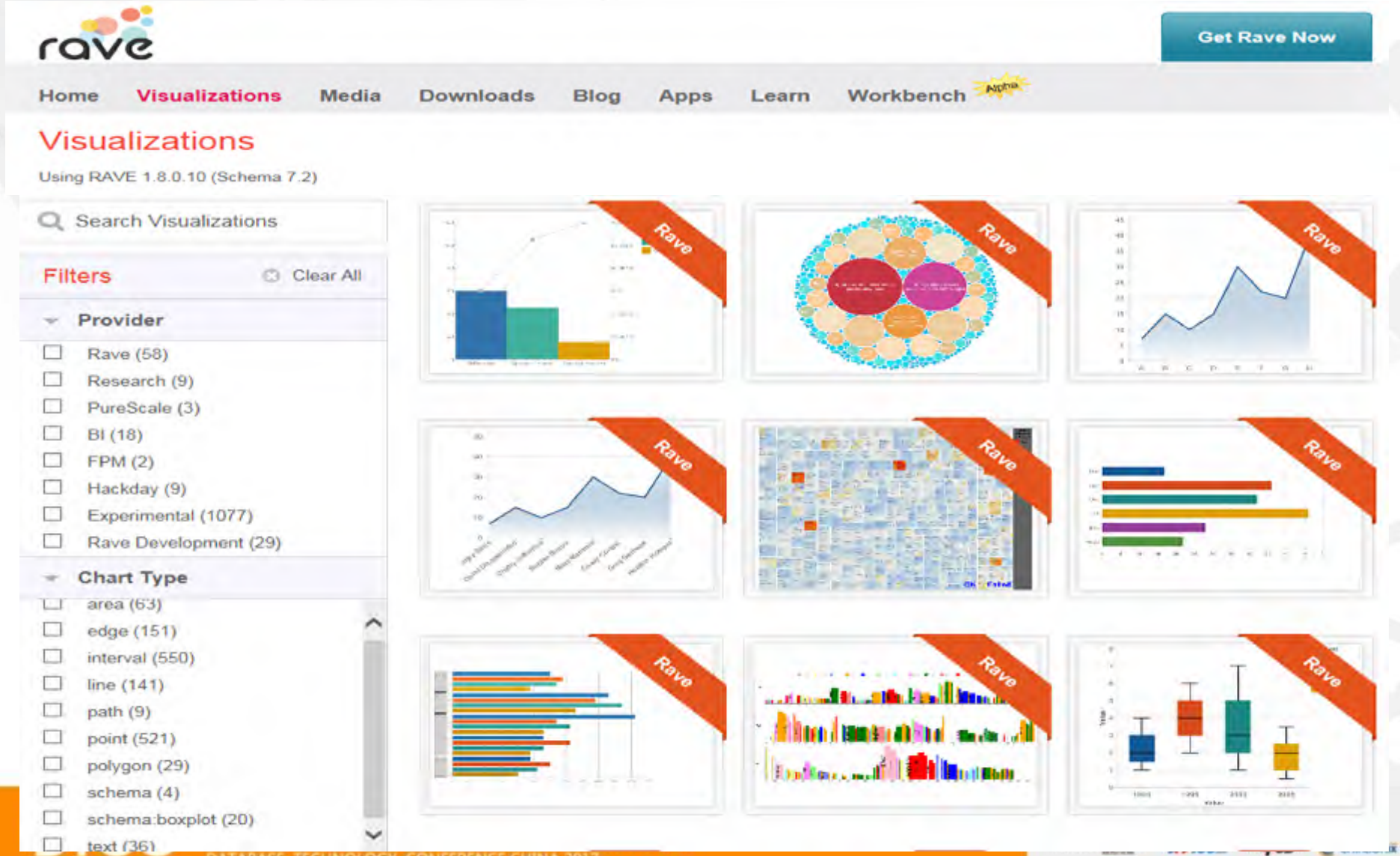
- 新的可视化需求随着大数据的出现在不断增加
- 每个需求都是特别的,难以直接满足

解决方案:

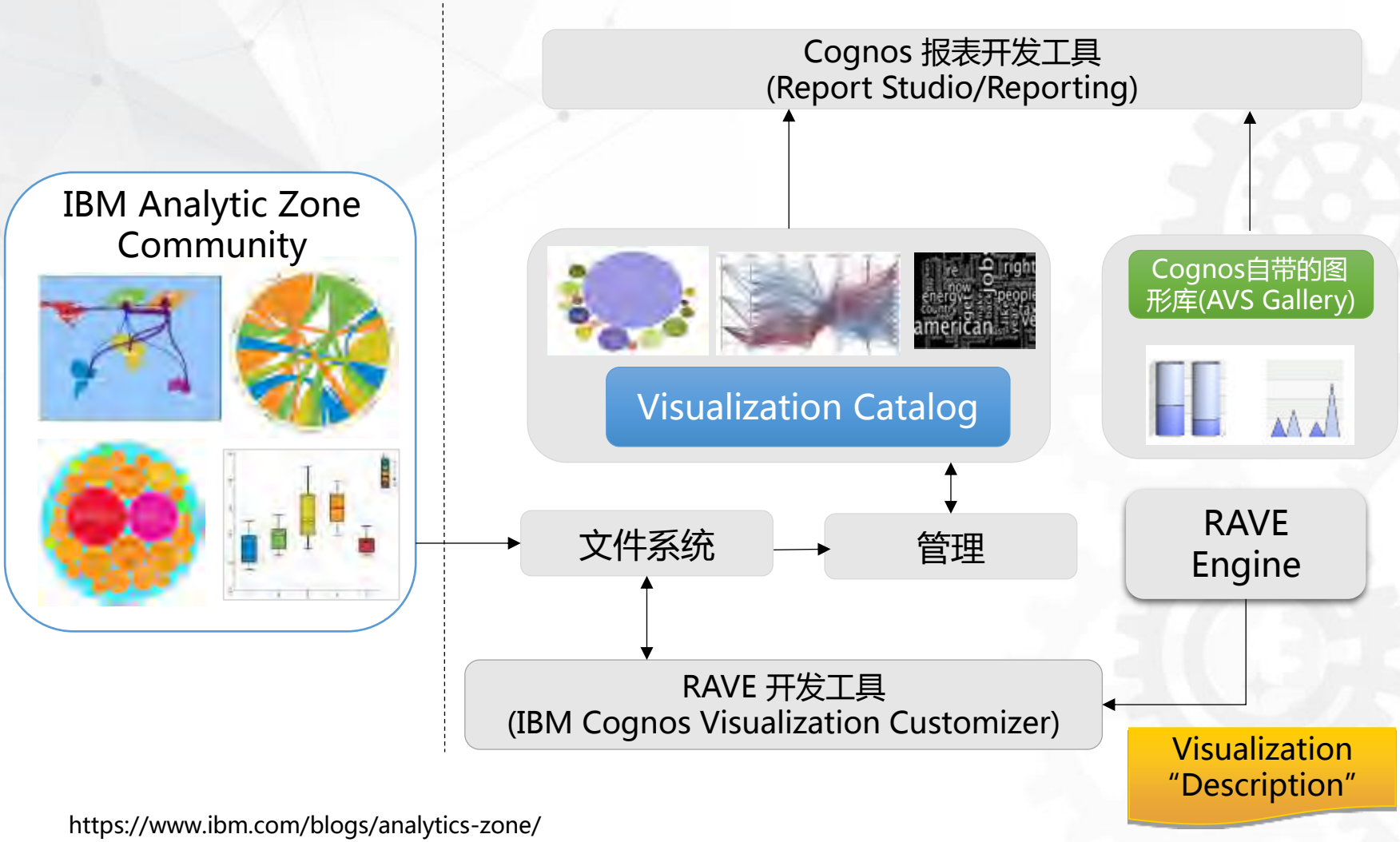
- 提供大量的可视化种类，也就是意味着可以方便地定制.



IBM RAVE图形引擎提供了大量的可视化示例

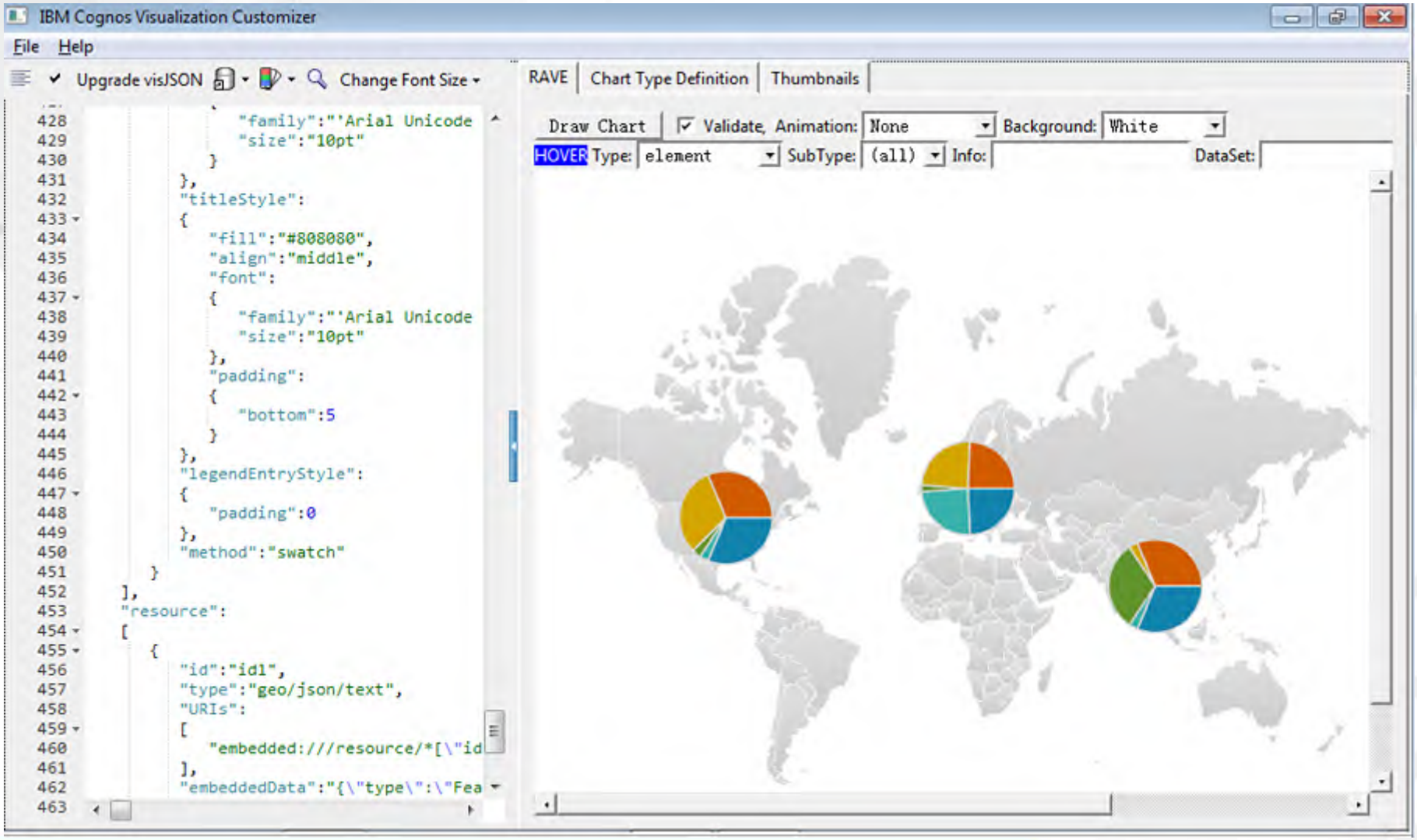


RAVE图形引擎与Cognos Analytics

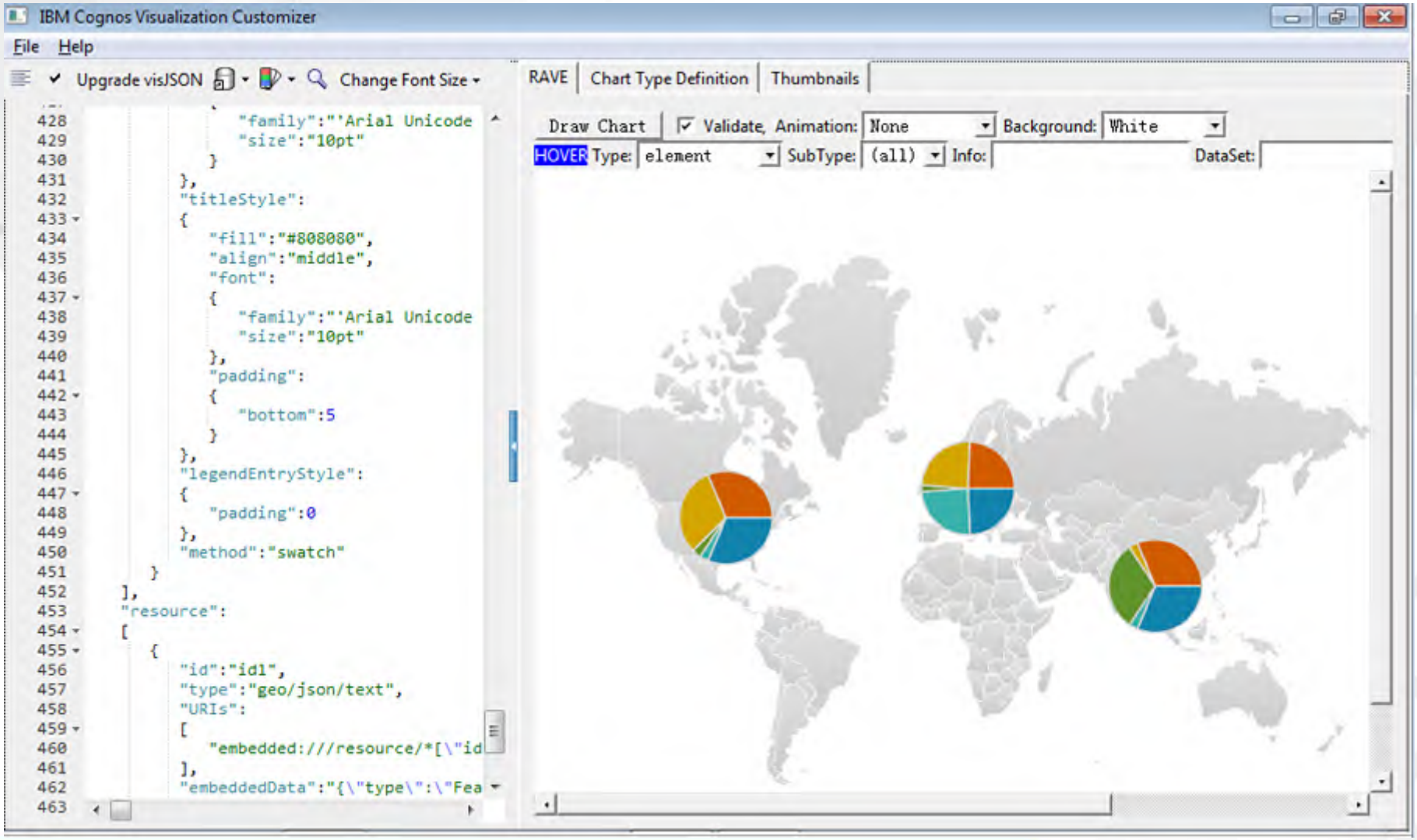


https://www.ibm.com/blogs/analytics-zone/

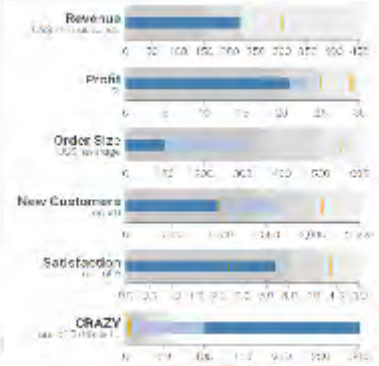
RAVE 开发工具
- IBM Cognos Visualization Customizer

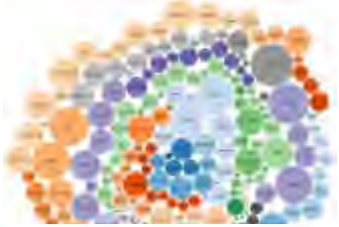


RAVE 开发工具
- IBM Cognos Visualization Customizer



RAVE 2更加值得期待...





用户交互界面 (UI)

简洁的用户界面能够满足用户各种分析需求：

数据上传

智能数据探索

仪表盘

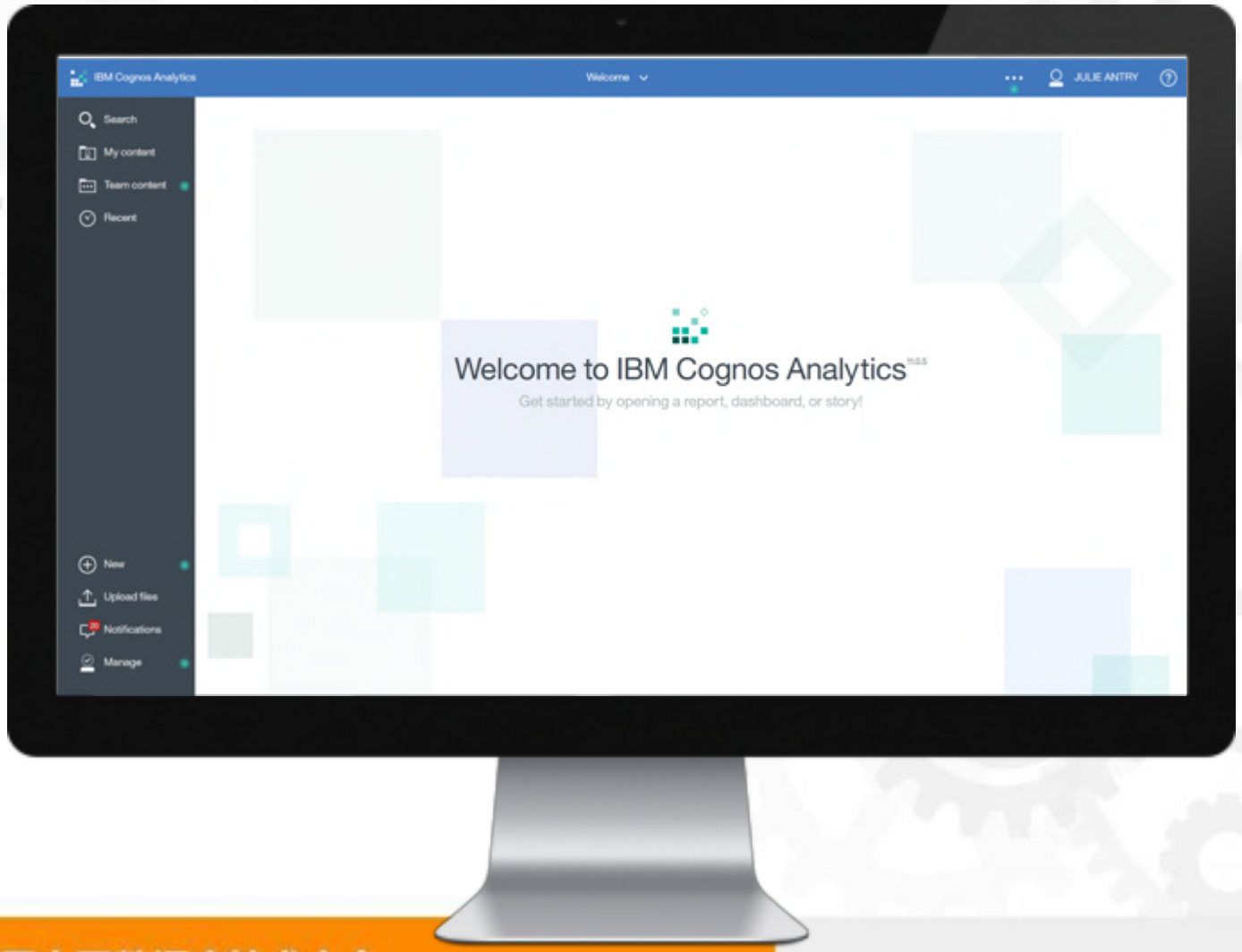
图表

创建故事

即席查询分析

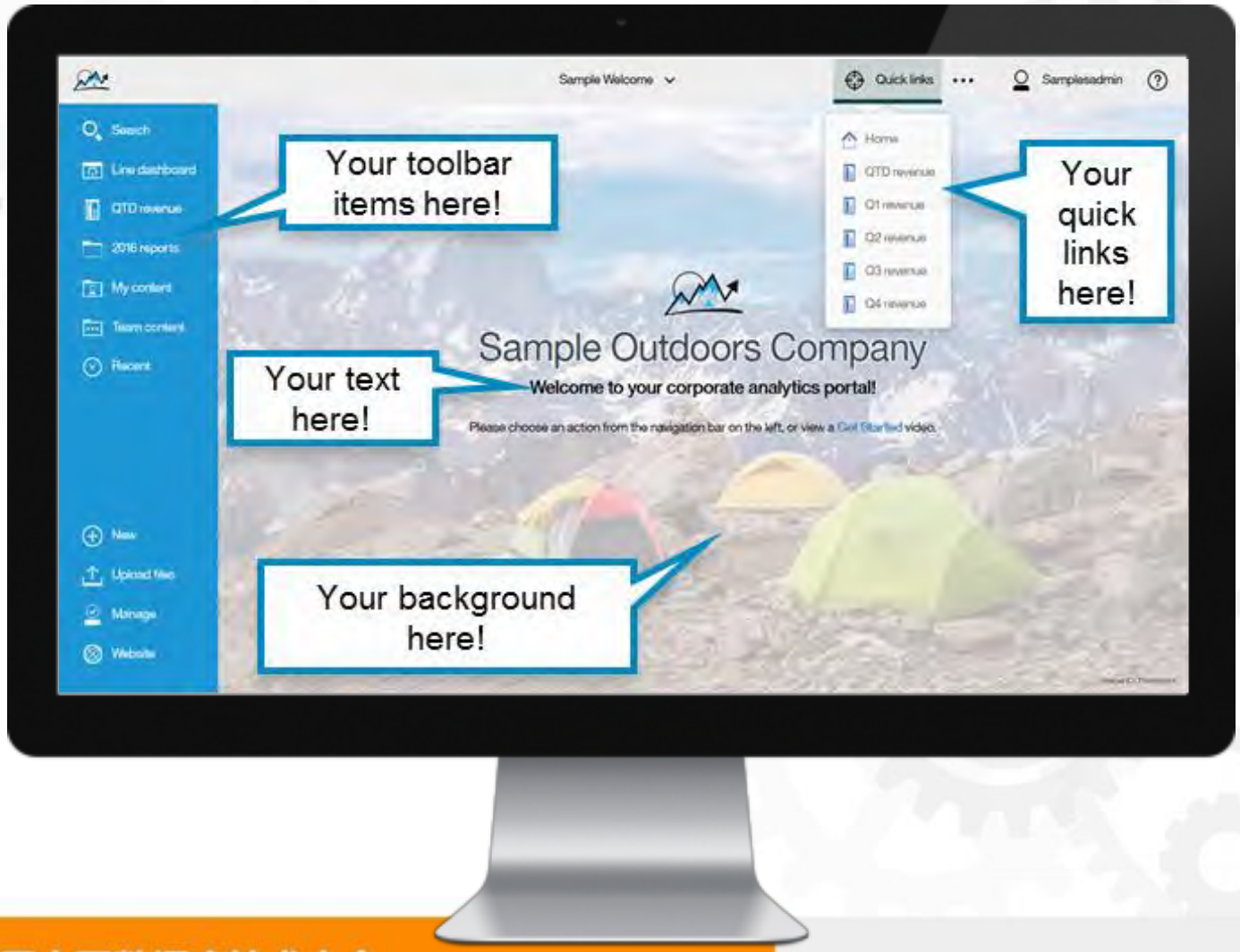
报表制作

数据建模



皮肤主题与界面定制化

主题允许自定义：
登录页面
欢迎页面
背景
导航列表项
快捷菜单
可以应用到用户组级别

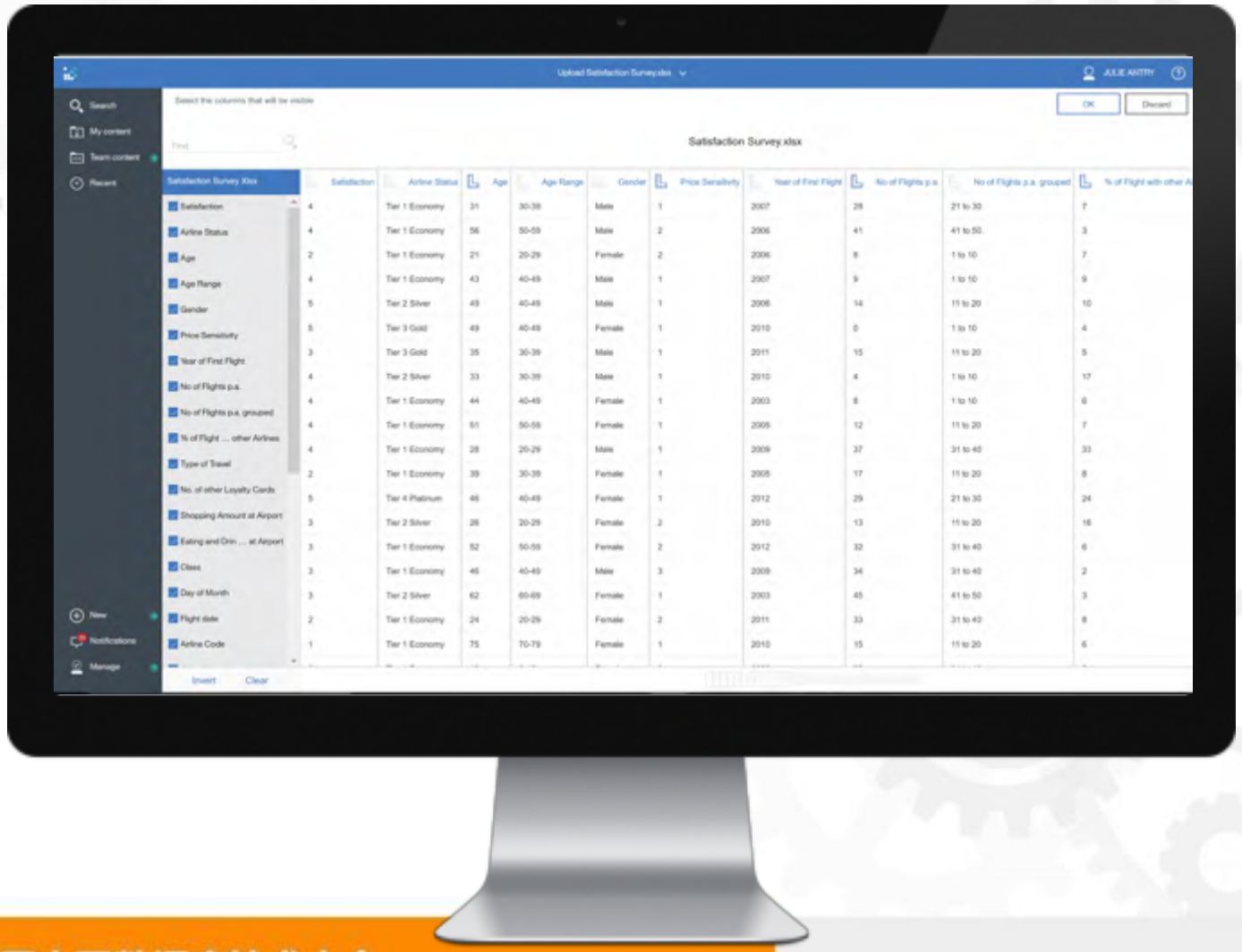


一流的管理报表

可交互式、
完美像素、
按需运行的报表，
或者后台批量生成
在线或者离线报表，
并且不限制使用用
户数。

数据上传

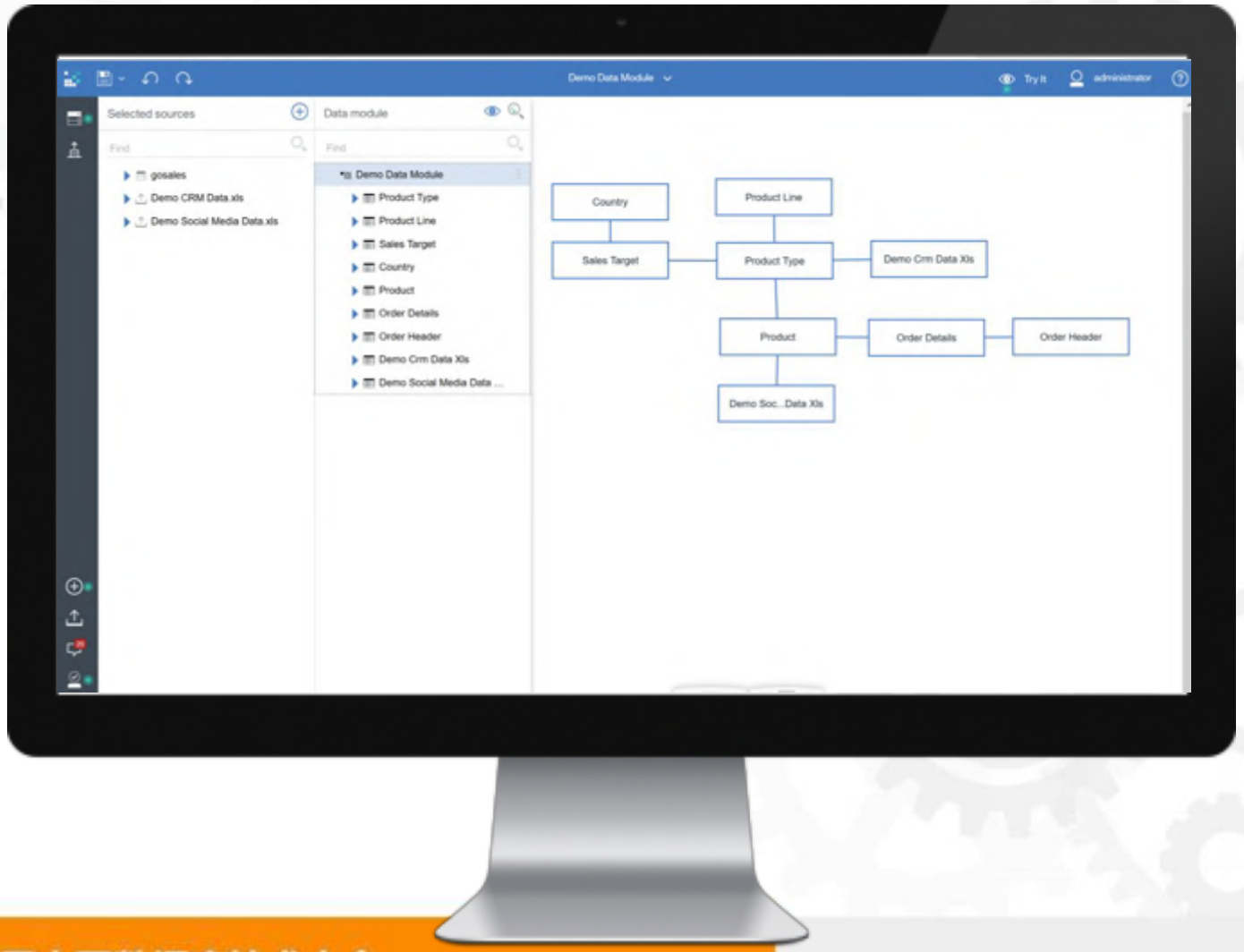
用户可以上传外部数据源，并且通过交互式操作立即对数据进行加工，并为即席查询、仪表盘提供数据，从而快速获得洞察。



数据建模

数据模块功能可以为用户随需进行数据混合操作。

用户可以将企业管控的数据与外部数据进行拼接。



高度可视化交互式仪表盘

Cognos Analytics的仪表盘是高度可视化，并且完全可交互的。

智能搜索与自然语言提供快速产生业务价值的智能数据探索。

可视化图形

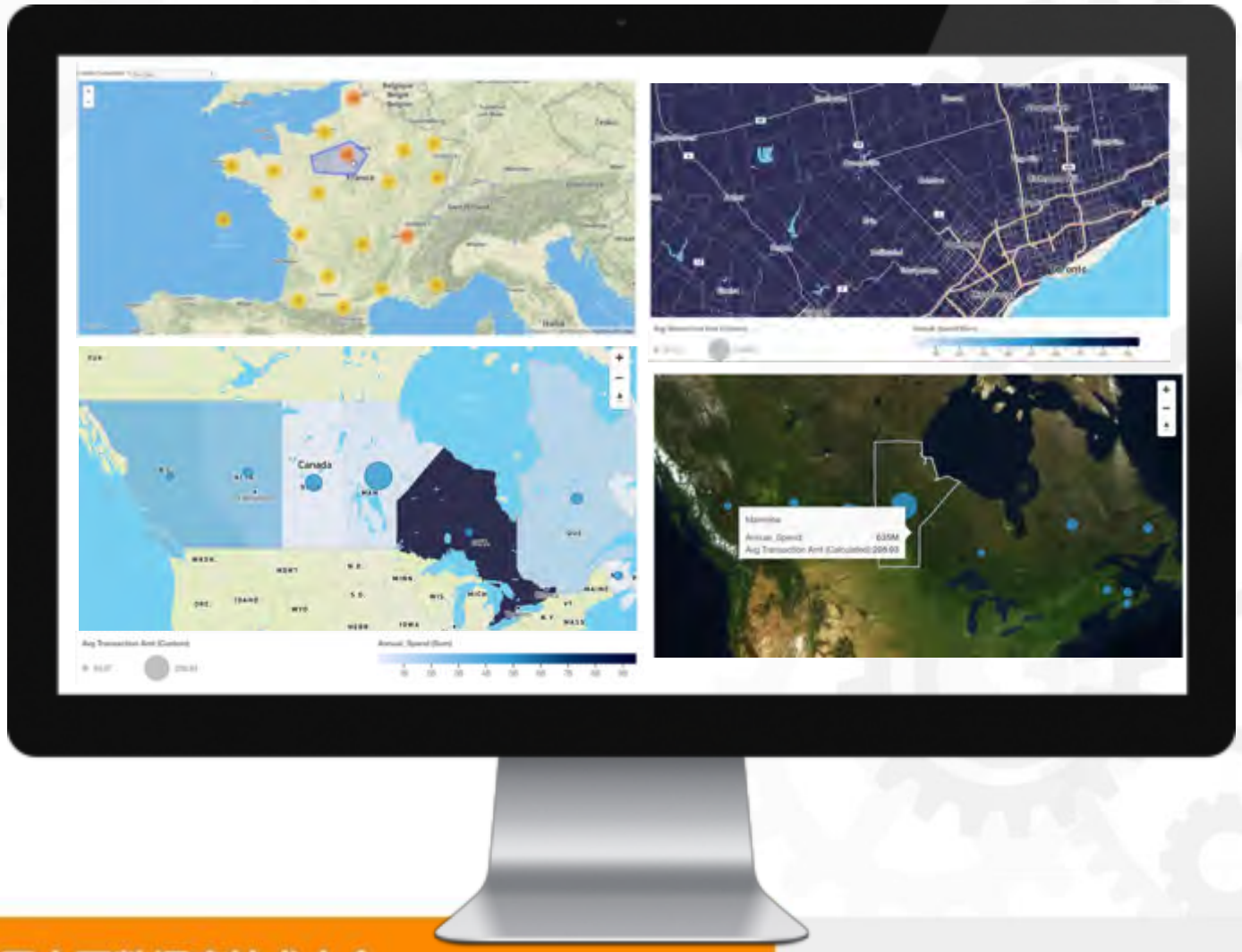
示例包括：

数据演示器（时间滑动器）
地理空间地图
信息图表
世界地图
热图
树图
网络图
散点图
箱图
龙卷风图
气泡图
雷达图
弦图
甘特图
面积图
仪表图

Cognos Analytics还可以提供更多的可视化图表，甚至合适图片、视频、网页链接等。

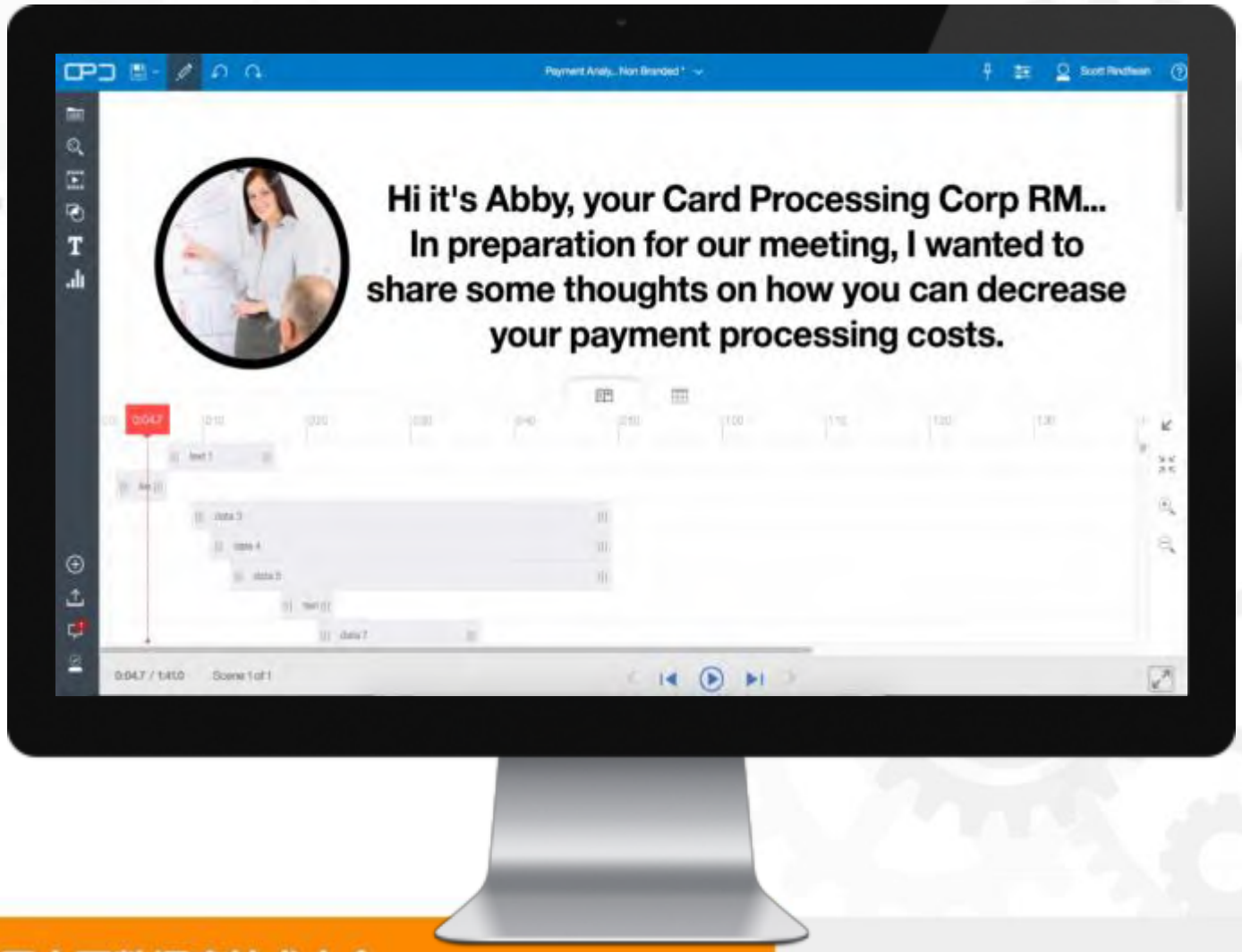
世界级的地理空间地图

Cognos Analytics与 MapBox合作，可以提供世界级的地理空间地图，包括街景视图与卫星照片视图。



自动演示分析

Cognos Analytics的故事与向导式的录制功能可以利用仪表盘的内容，增加动画、图片、文字等信息来完成最终的演示故事。



客户对IBM Cognos Analytics的反馈

> IBM Cognos Analytics combines self-service and managed data analytics capabilities to support new features, including intent-driven modeling, smarter search that works in-context and a single environment for all types of business reporting.

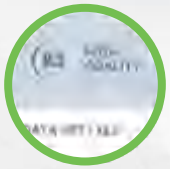
Charles King,
Industry analyst

> One of the most impressive features is the improvement and prominence of the search functionality and its harmonization with what has been described as the “intent-driven” experience of authoring and modeling.

Scott Misage,
Engagement Manager

Watson Analytics: 智能数据探索

Watson Analytics在业务用户的易用性和强大的功能之间需求平衡



智能元数据能够自动整理数据源中的数据以更好的用于分析

认知指引在分析的开始给予分析建议，并提供自然语言对话方式进行分析

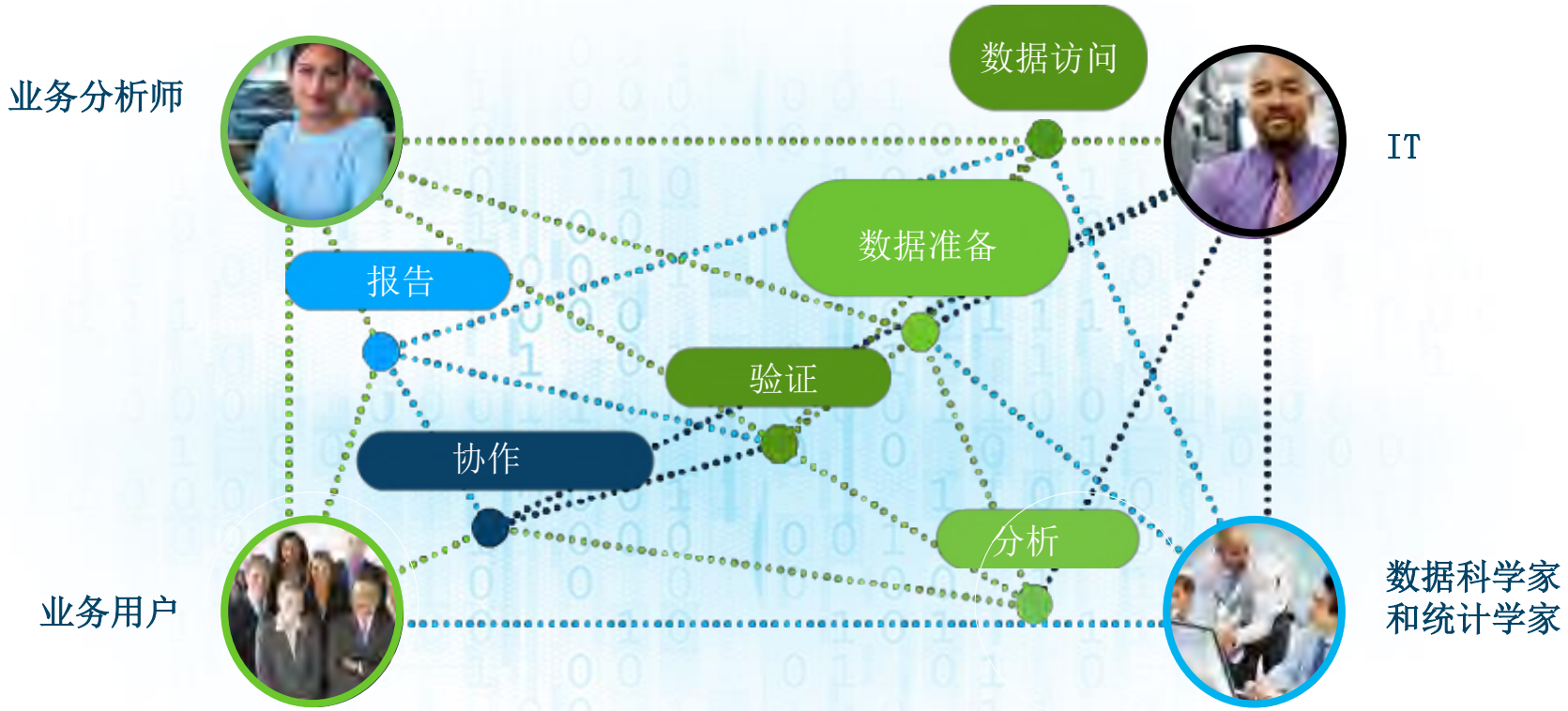
高级分析不需要专业知识就可以发现数据模式和内在关系

访问**社交媒体**数据，并对其进行分析,从而获取外部客户的观点和态度

以**仪表板**的形式可视化支持您的决策，传达决策结果

即使是简单的分析项目，也需要繁杂的步骤和大量的人员

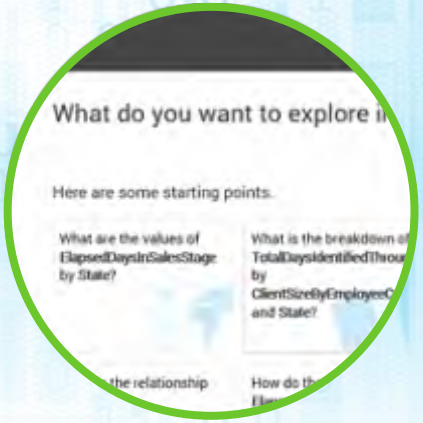
过程往往错综复杂



IBM Watson Analytics

云端自助式分析功能

单一分析体验



完全自动化智能

自然语言对话

引导式分析发现

Watson Analytics客户成功案例

Honest Café – Brewing up the Perfect Business Model Based on Customer Insight

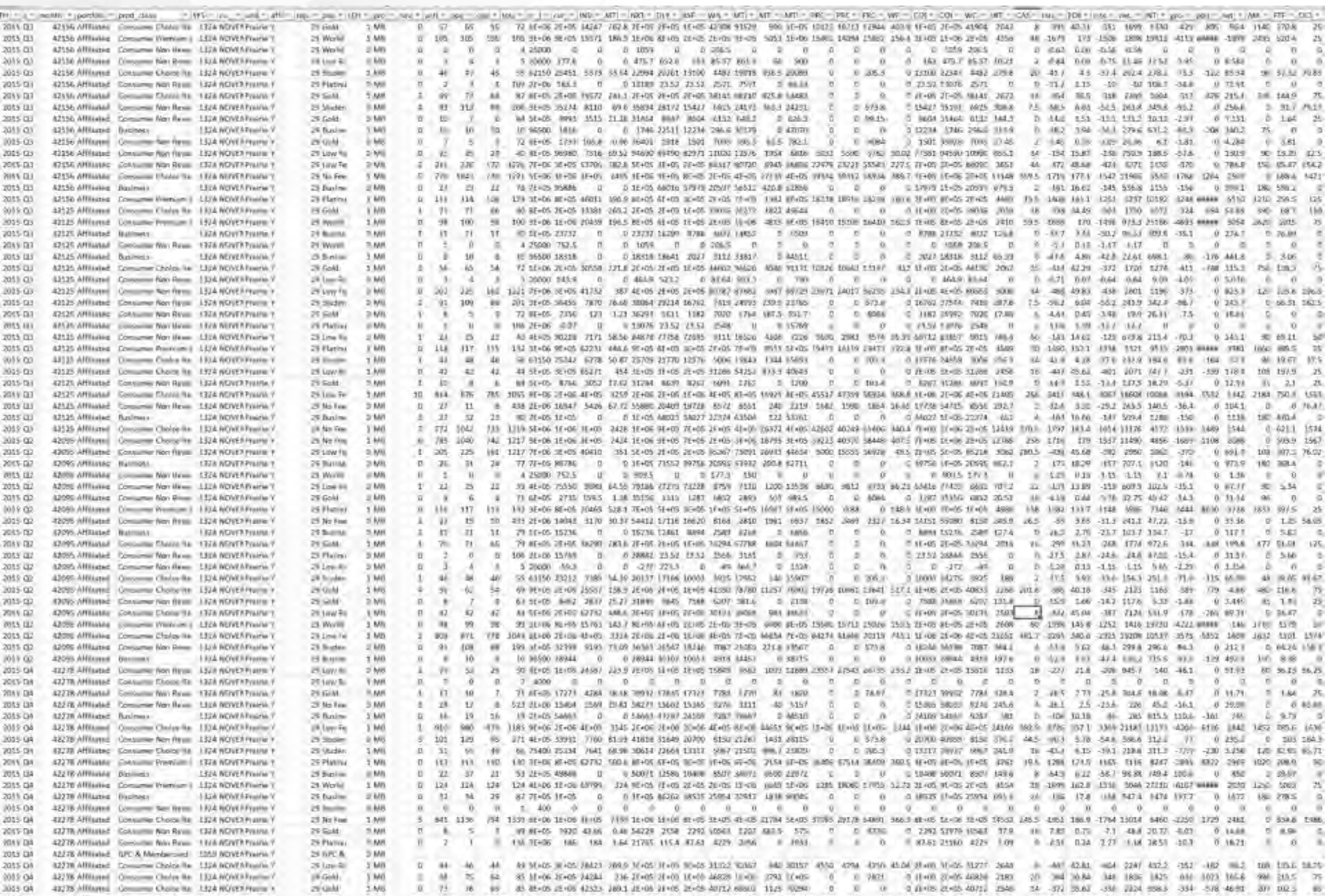
The Transformation

Vending specialist Honest Café offers premium-quality coffee and snacks without the staffing costs of a traditional café. Cognitive analytics revealed that customers see a visit to Honest Café as a social experience – leading the company to redesign its retail spaces, adding seating areas to encourage longer visits and upsell.

"Cognitive analytics has transformed our understanding of customer behaviour – and our retail strategy."

- Mark Summerill, Head of Product Development, Honest Café

这是您的数据...



这是在IBM Cognos Analytics里的数据

Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

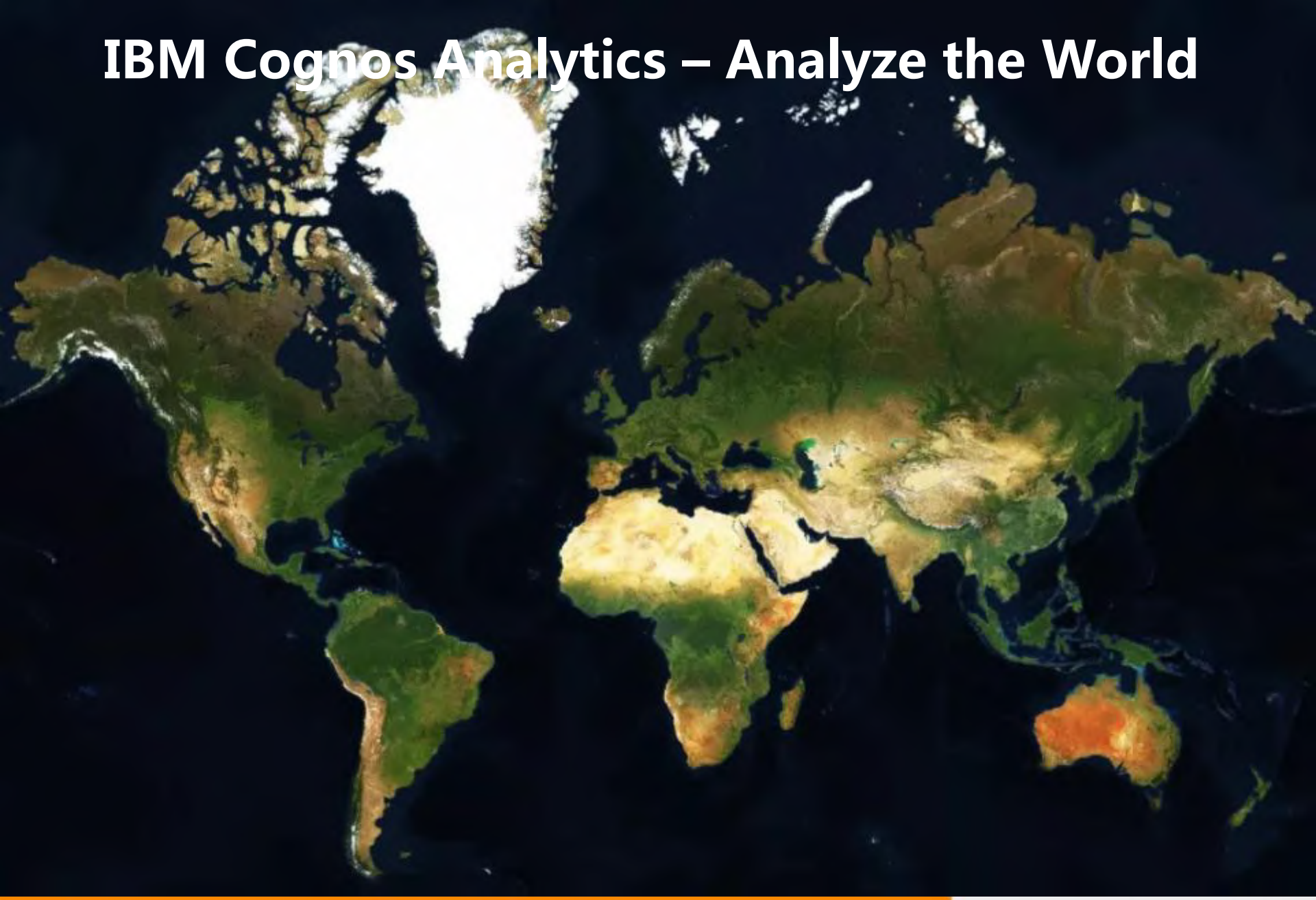
Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

Cognos Analytics: Art of the Possible

IBM Cognos Analytics – Analyze the World



DB2 11 and Cognos Analytics 11 - Fast on Fast

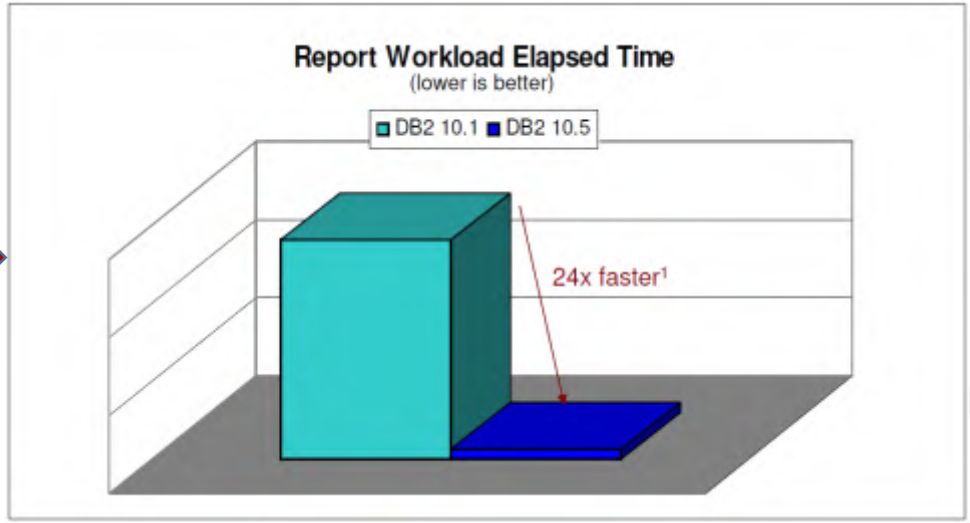
数据仓库类分析:
用途最广、扩展最强的内存计算数据仓库平台

大规模扩展数据仓库，内存计算的性能

MPP BLU 扩展能力
- PB 级内存计算数据仓库

下一代内存计算性能, 功能和负载

- **更快的ELT/ETL 性能**
- **优化了更多的查询负载**
- **支持了更多的功能**
  - 派生列
  - RCAC
  - OLAP + BLU性能



关注微信公众号，获得大数据分析最新最热咨讯

THANKS

SequeMedia 盛拓传媒 IT168.com ITPUB ChinaUnix