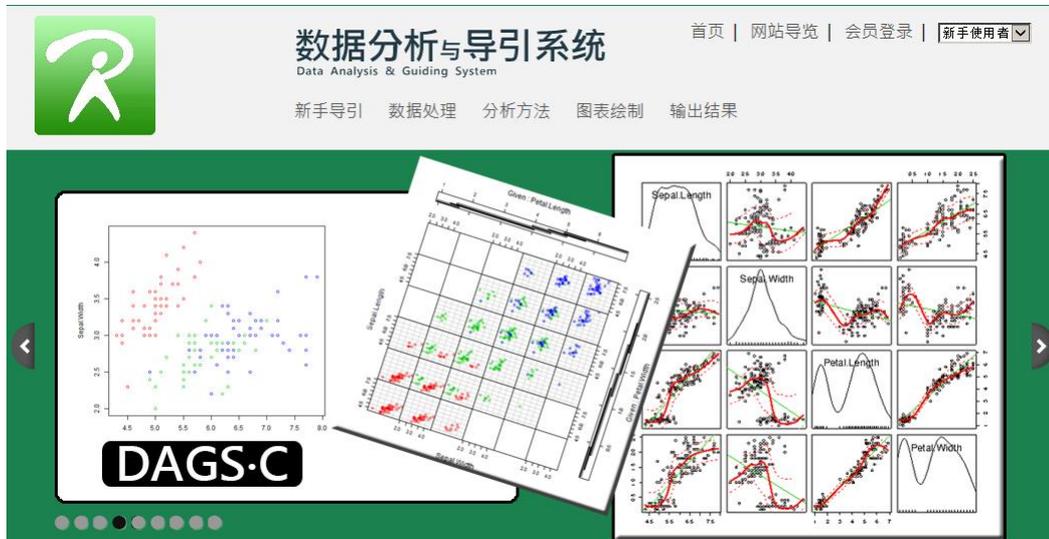


第一个由华人统计学家基于R语言研发的数据分析云平台

# R-web : 大型多人在线数据分析协作平台



林祯舜 博士  
先锋信息技术有限公司  
吉林大学/兰州商学院

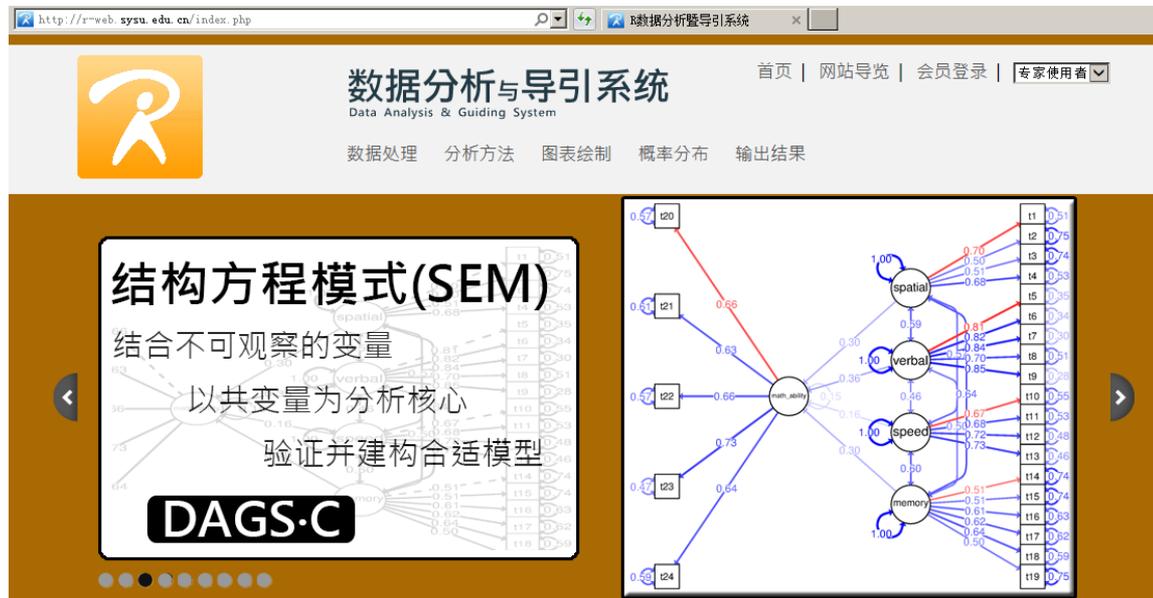
天津

2014.12.21

# 大纲

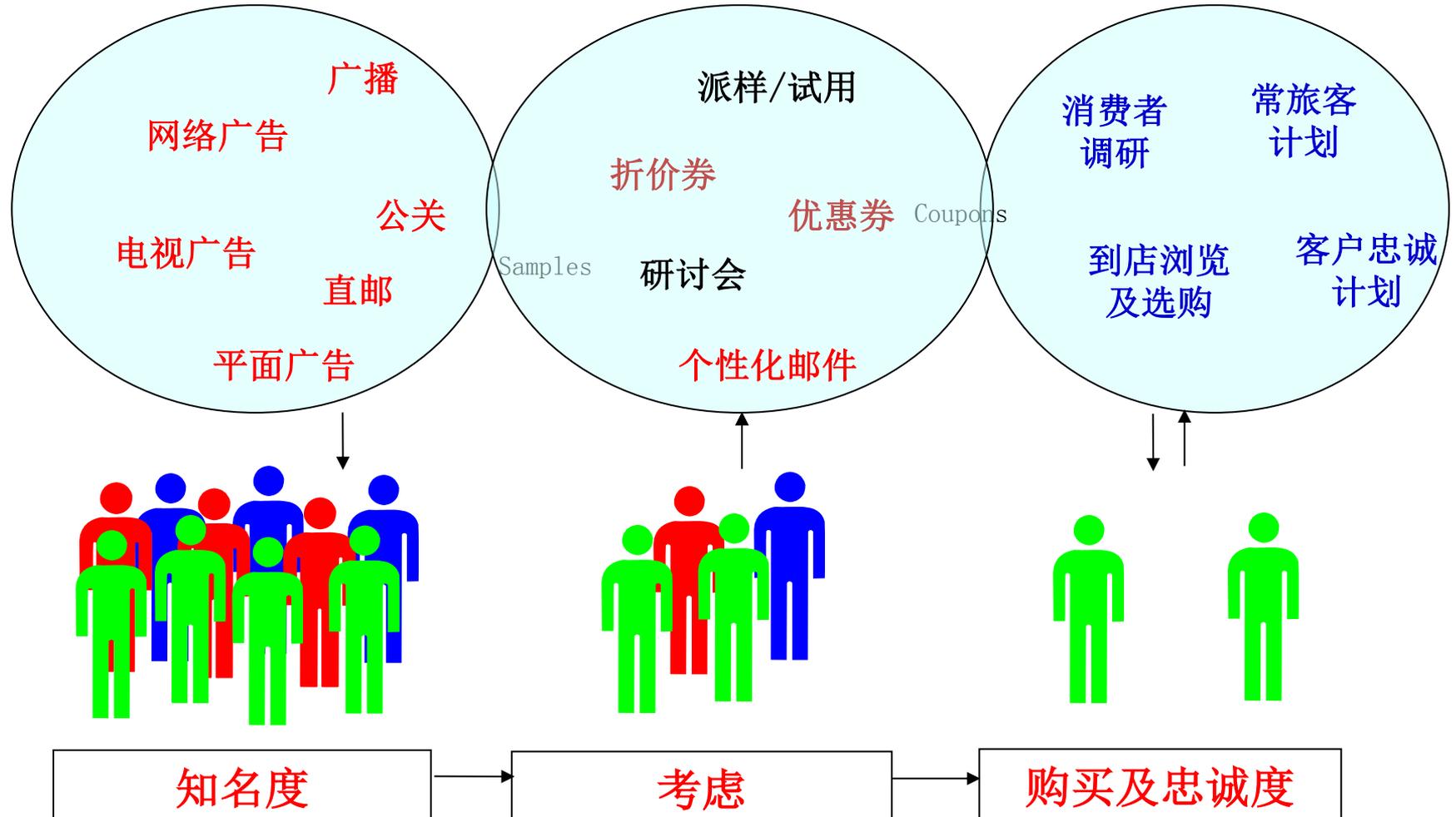
- 数据来源的根本性变化-  
以营销的演进举例
- 分析的组成要素
- 平台设计理念
- 平台系统演示
- Q&A

# 数据来源的根本性变化 -以营销的演进举例



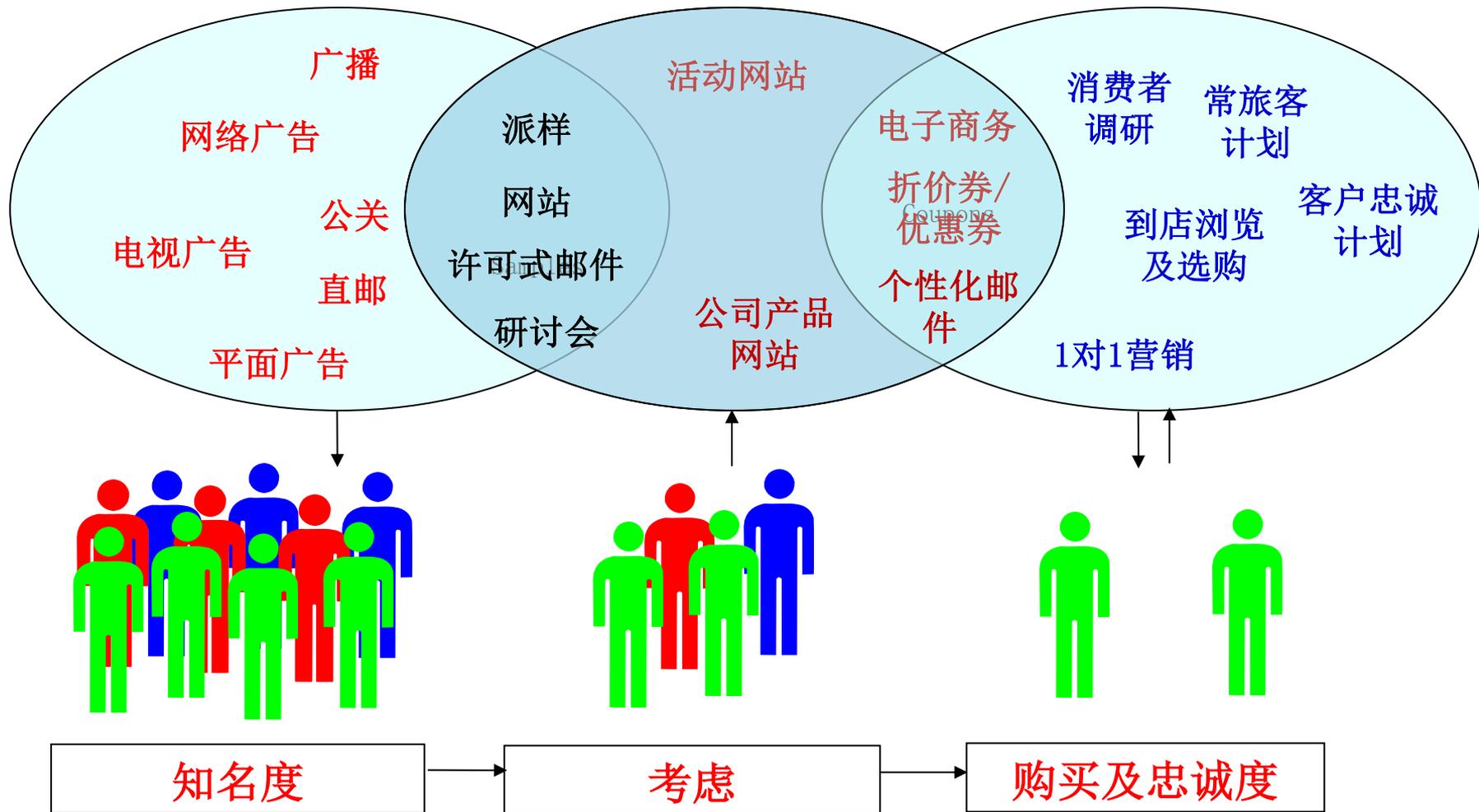
# 三十年前的营销 – IMC

不同阶段的边界清楚



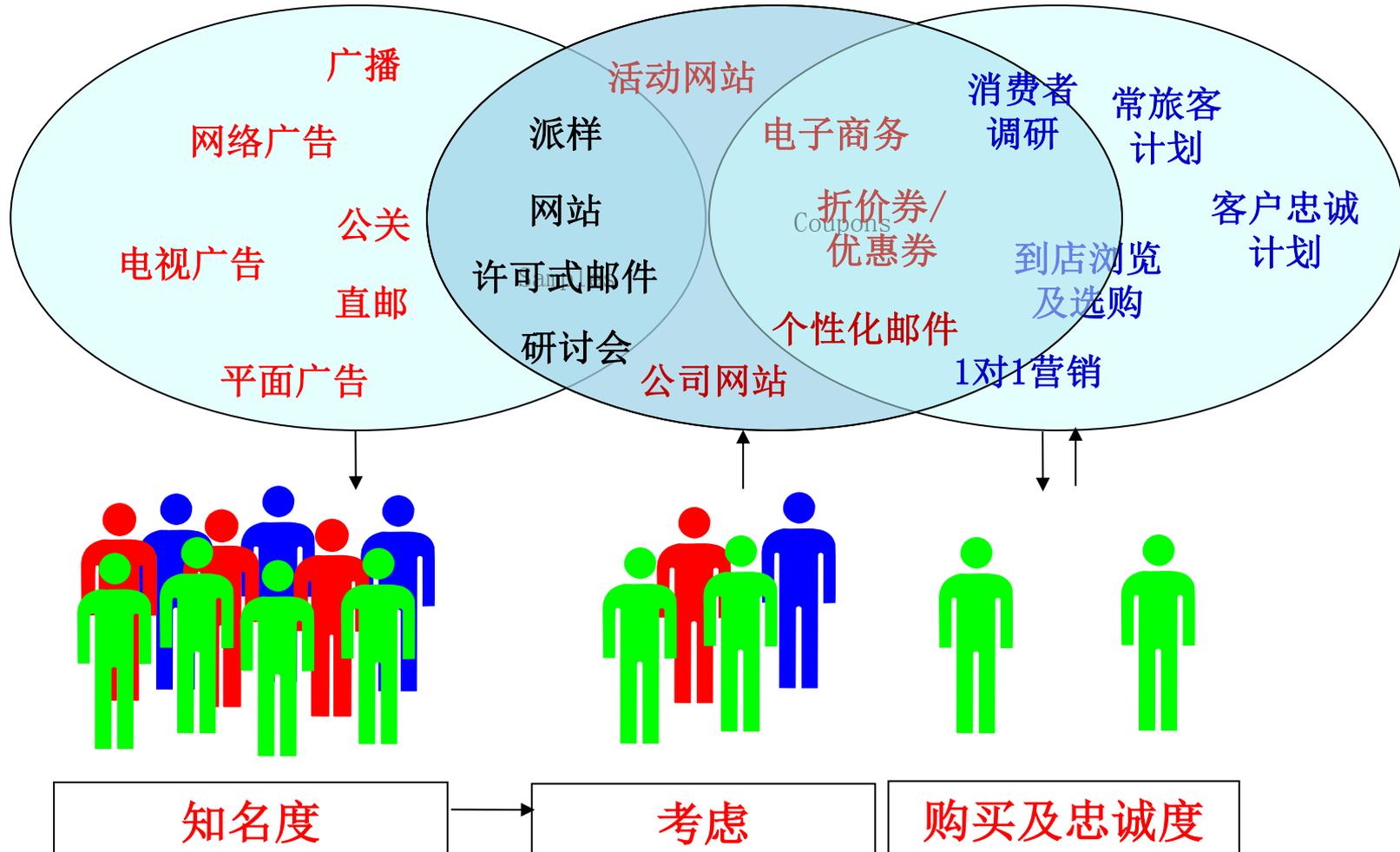
# 十五年前的营销 - CRM

信息化和互联网让边界逐渐模糊



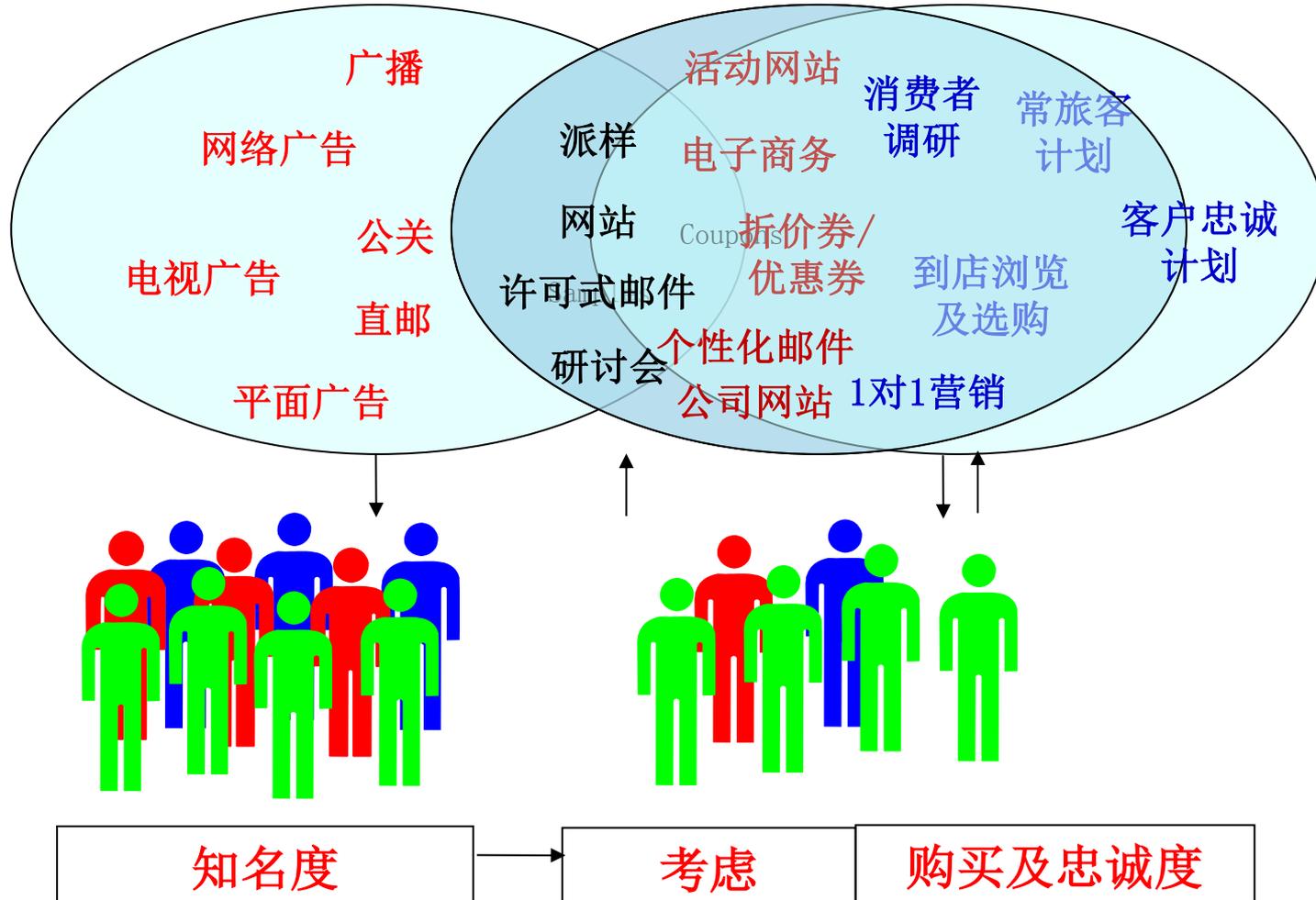
# 五年前的营销- Digital

互联网和电子商务让边界逐渐融合

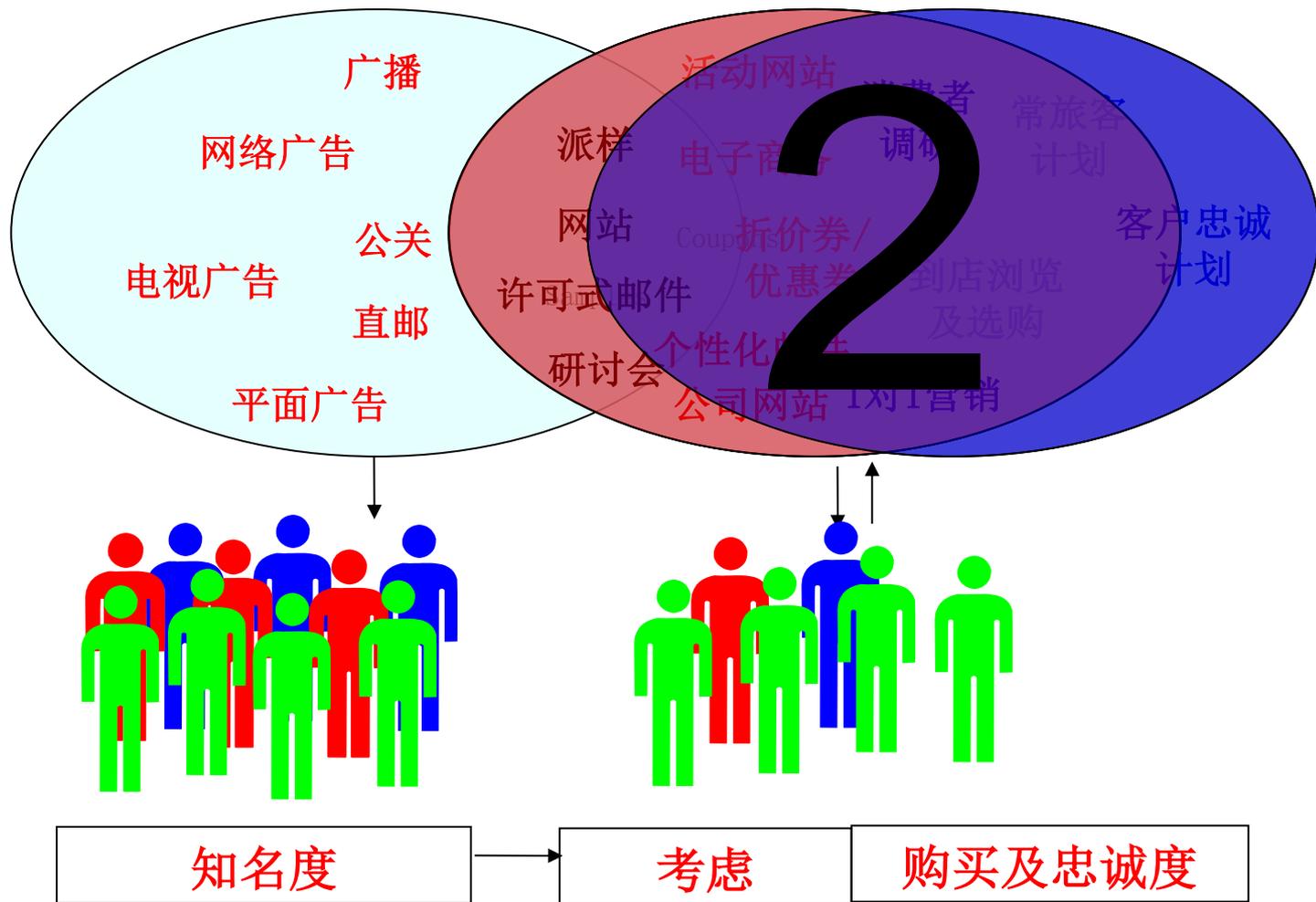


# 现在的营销

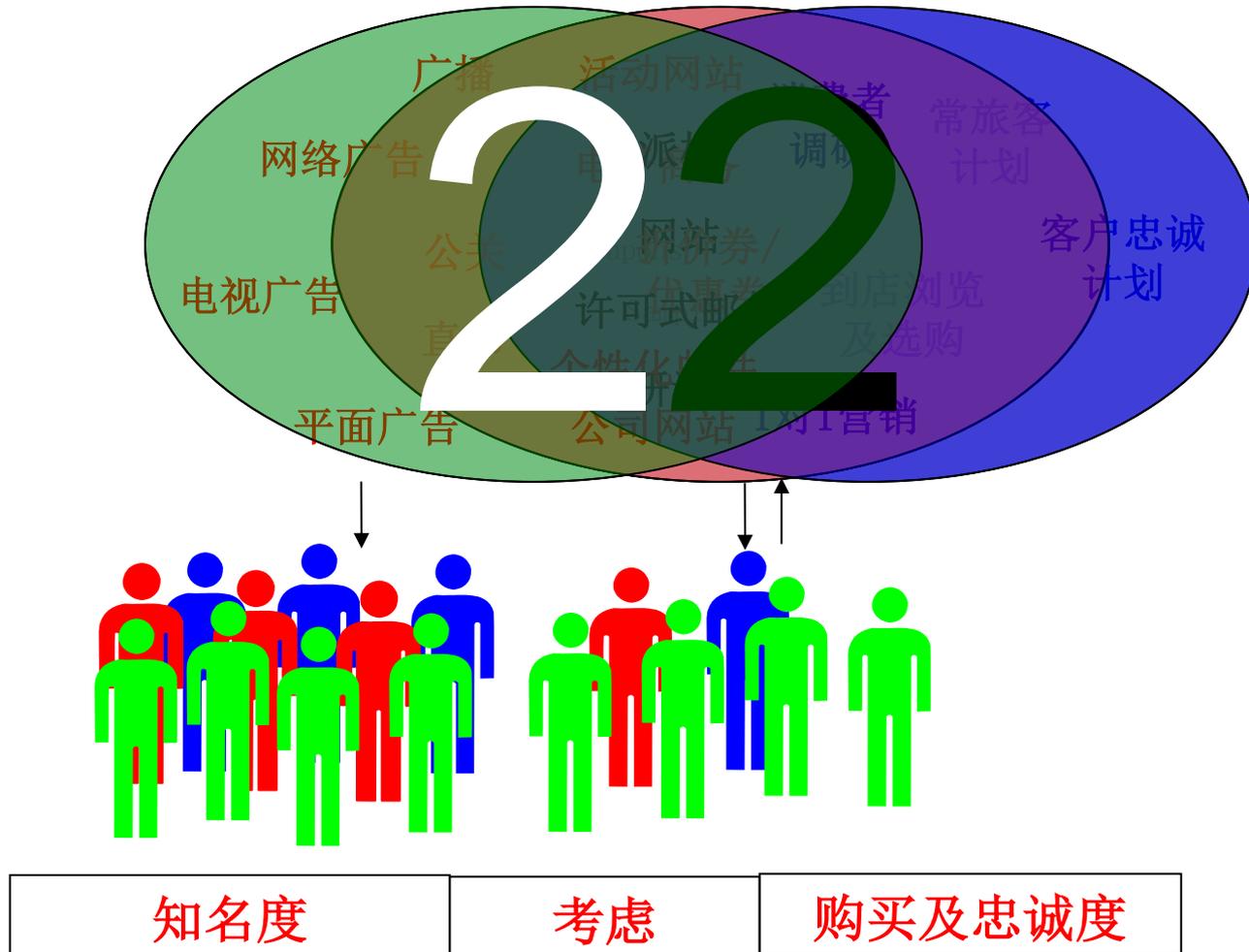
## 跨界



# 现在的营销-O2O



# 未来的营销 - ~~2020~~



# 未来是大数据驱动的营销

营销的**本质**没改变，因为**技术**和外环境的进步而产生了**典范的转移**。在这个典范的转移过程中，催生了许多新观念，在中国大地上，**O2O**火了!!

其实，从**学术**的角度而言，营销学者并不常用**O2O**这个概念，而是用**全渠道 (Omni-Channel)**零售或是**多渠道 (Multi-Channel)**零售这些名词。

# 未来是大数据驱动营销

不管你用那个概念，真正的核心是

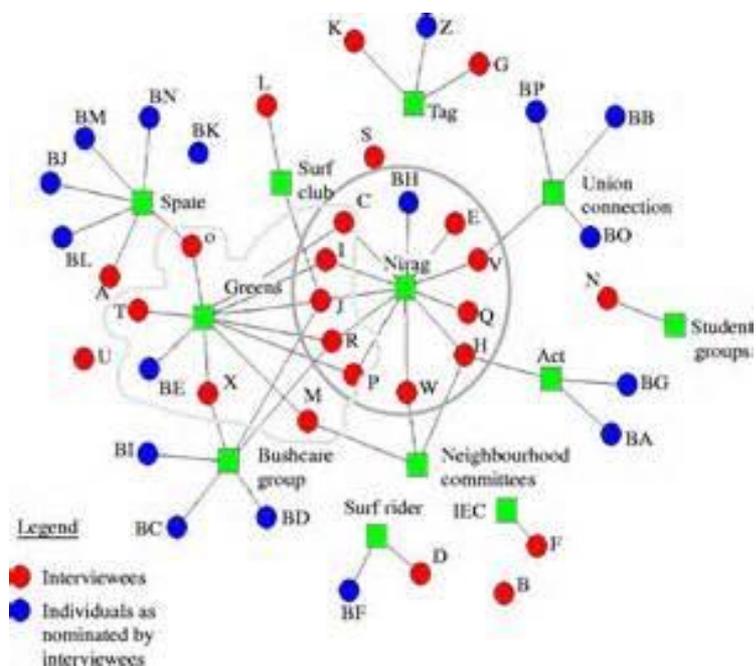
## 数据

只有在业务的环节建立了闭环，数据才有多样性，当数据的外部性逐步体现，才有大数据应用的基础，有了大数据，才能开始真正建立我们的商业模式，那个时候才是O2O真正的核心价值---完善业务生态体系。



# 数据来源已发生根本性的变化

## 核心的观念：镶嵌、连接



## 交叉复用、全息可见

# 模型的组成

# 模型的组成有哪些？

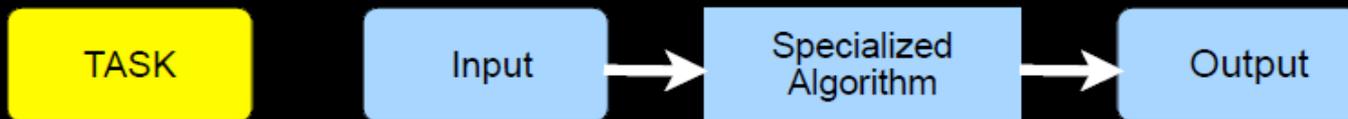
学数学/算法(理论)的人这么说

# 模型的组成有哪些？

## Our work today ...

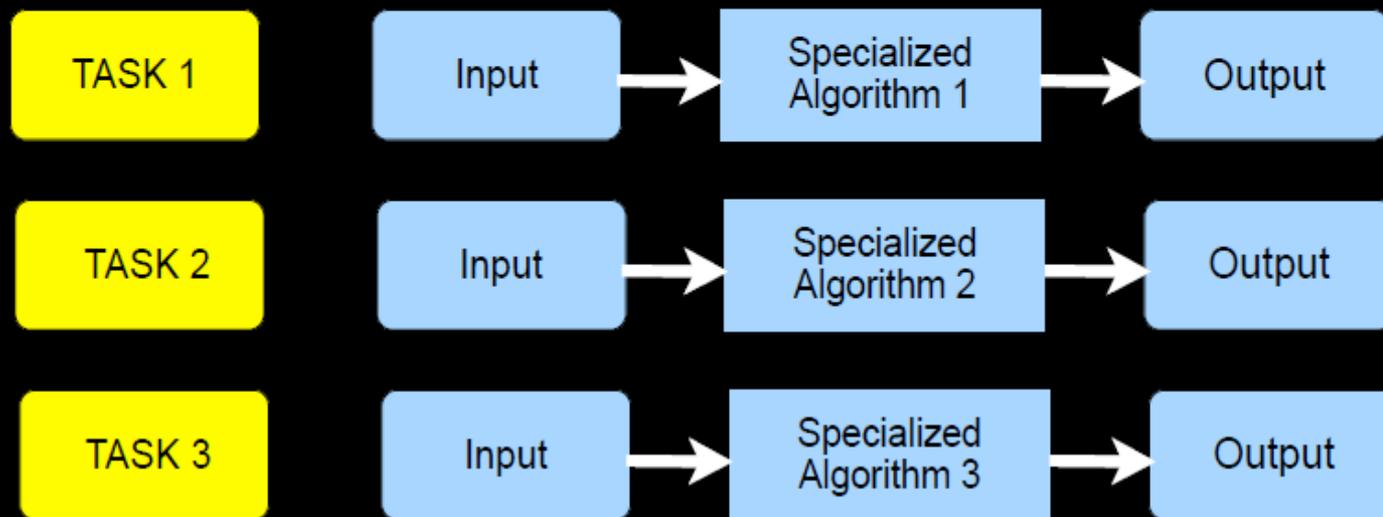
We typically ...

1. Formalize learning / mining task
2. Design algorithm / technique to use
3. Implement the algorithm
4. Use and distribute the software



# 模型的组成有哪些？

And do it again ...



# 模型的组成有哪些？

## Our work today ...

We typically ...

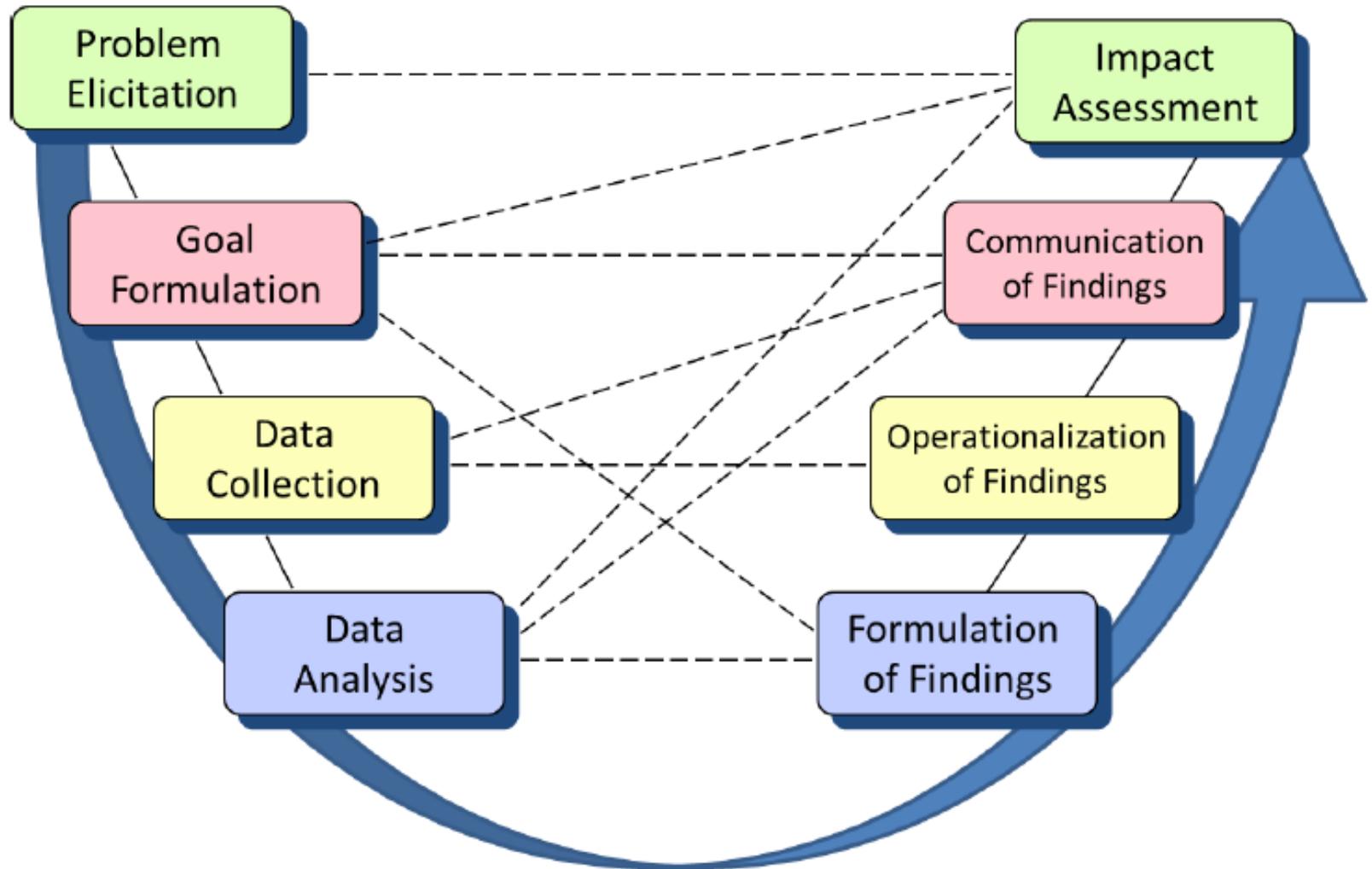
1. Formalize learning / mining task
2. Design algorithm / technique to use **hard**
3. Implement the algorithm
4. Use and distribute the software



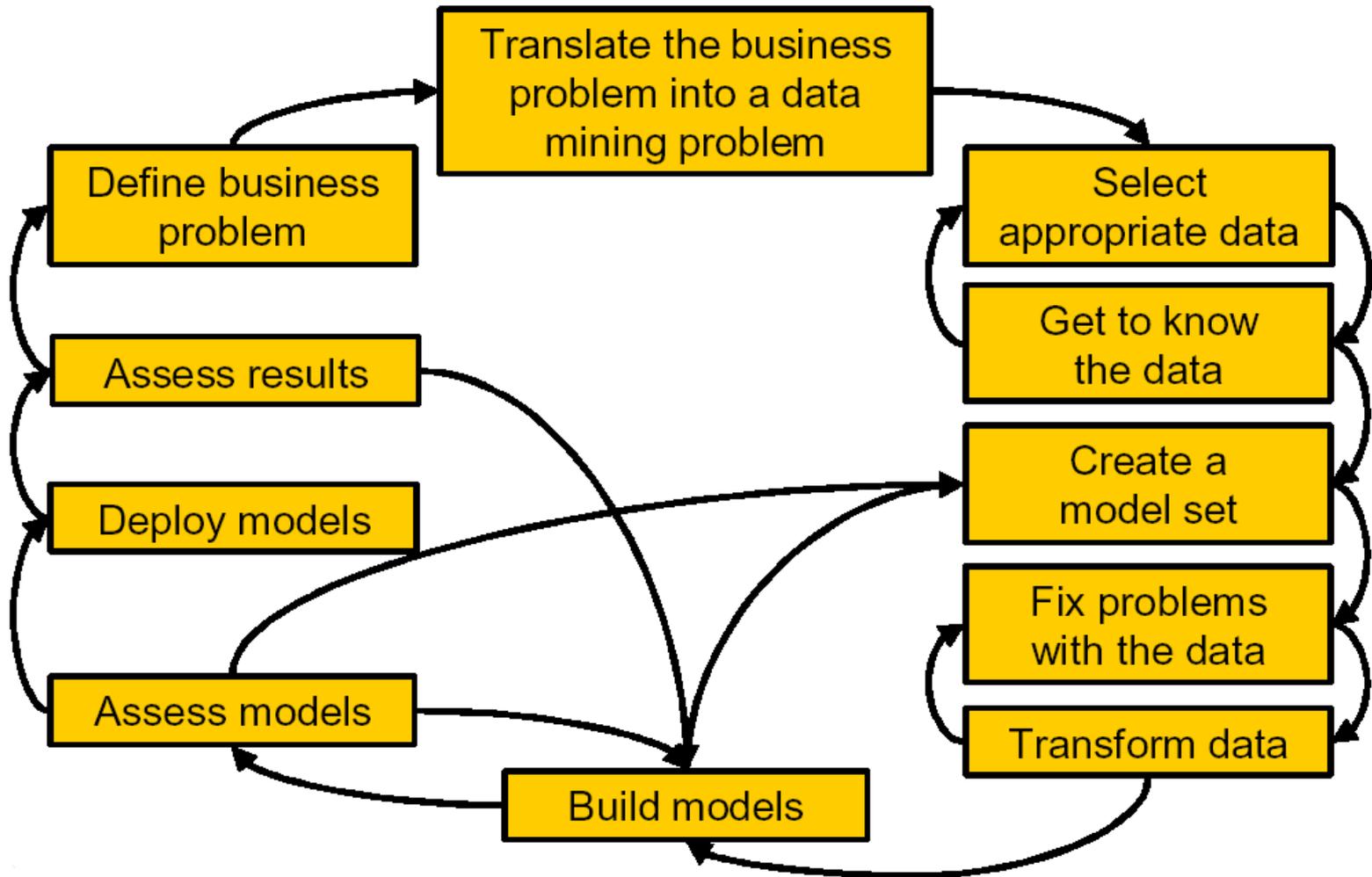
# 模型的组成有哪些？

学统计/分析(玩数据)的人这么说

# Statistics: A Life Cycle View



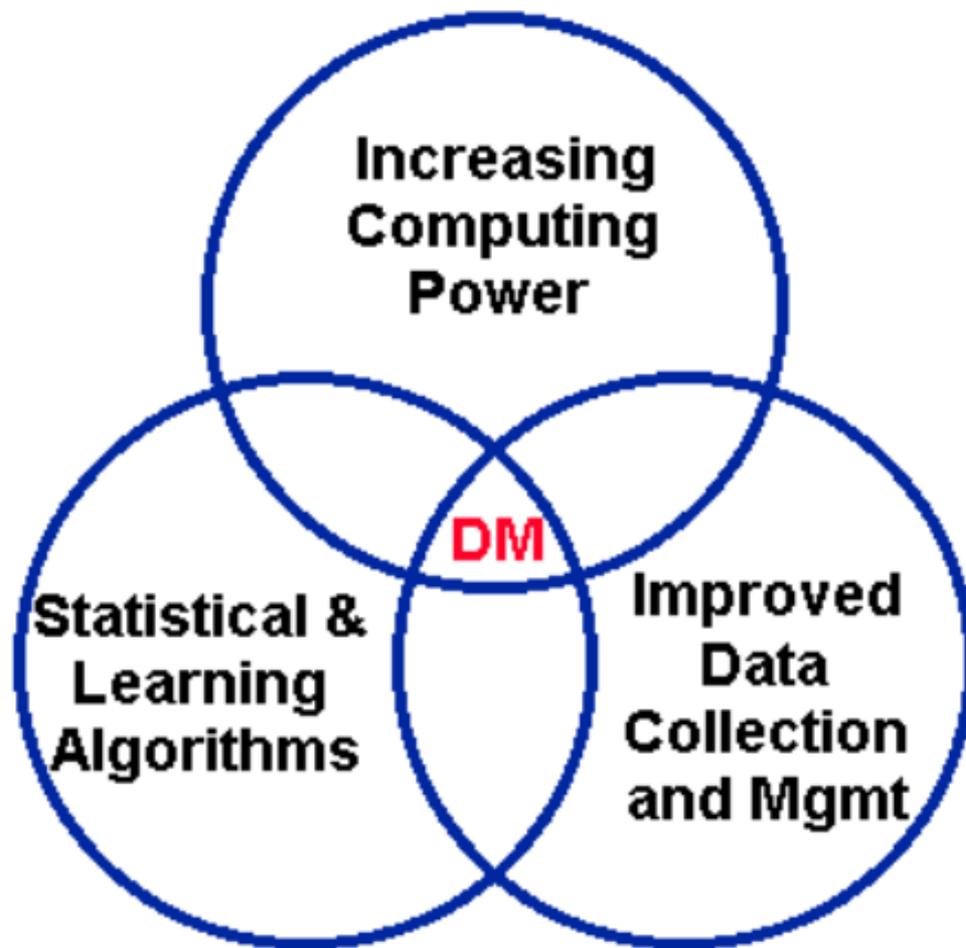
# Data Mining Is Not a Linear Process



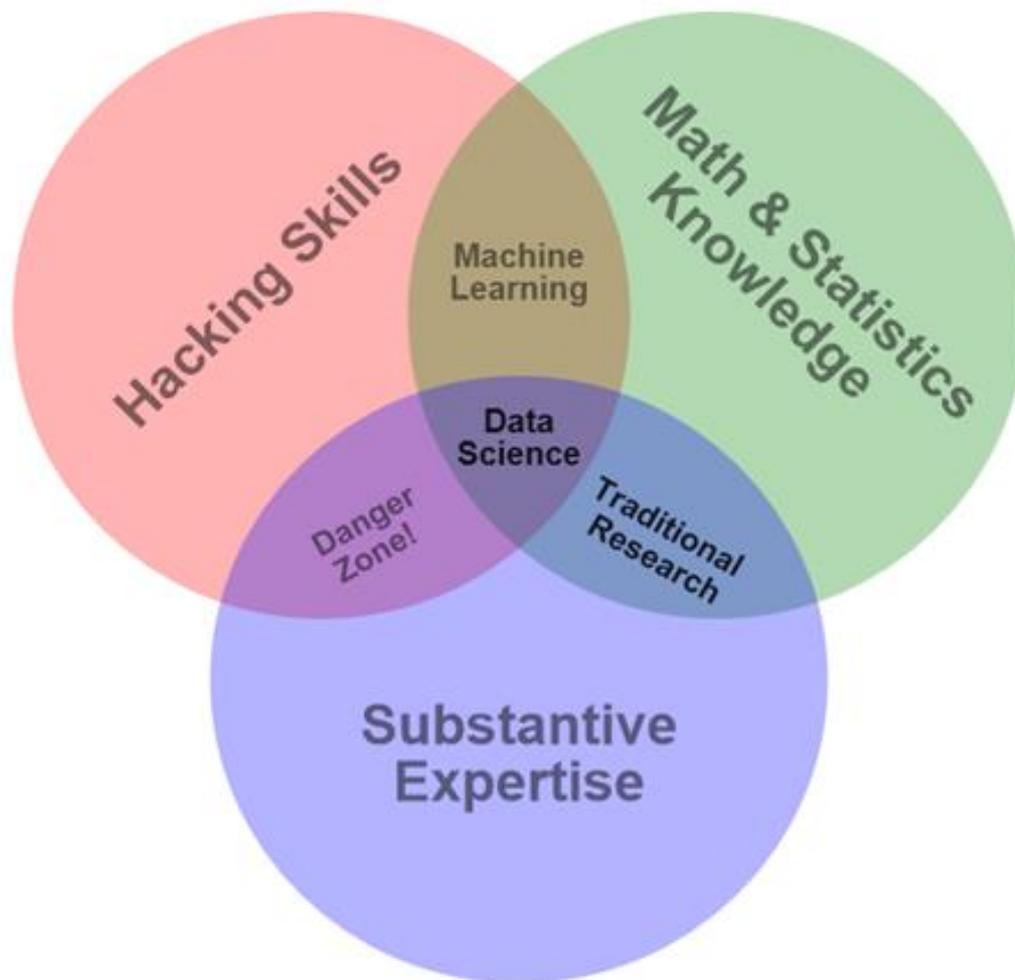
# 模型的组成有哪些？

学商学(搞管理)的人这么说

# 模型的组成有哪些？



# 模型的组成有哪些？



# 模型的组成有哪些？

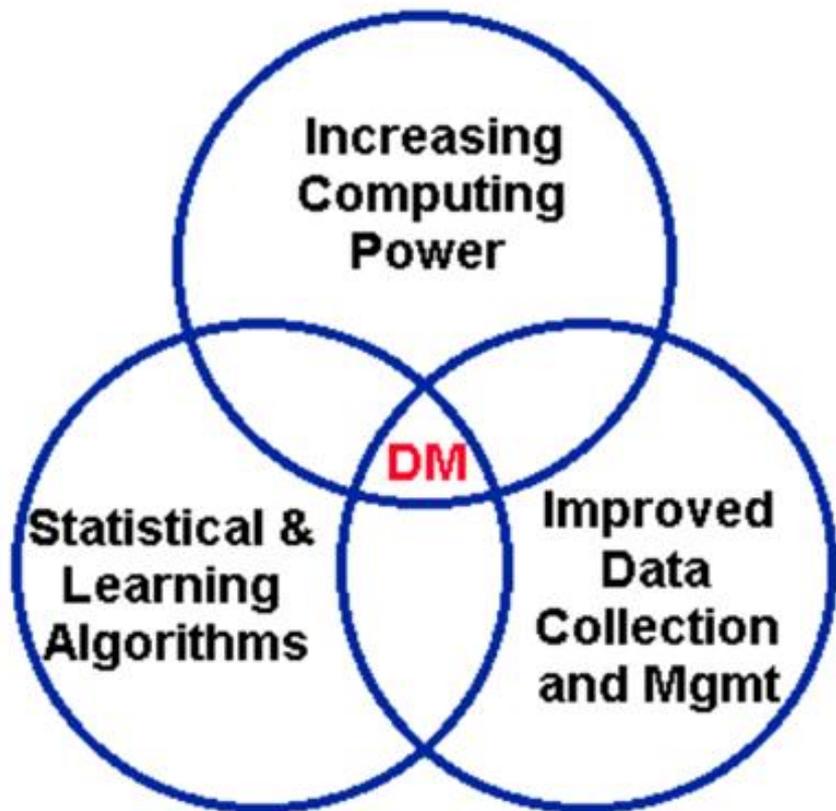
学哲学的人这么说

# 模型的组成有哪些？

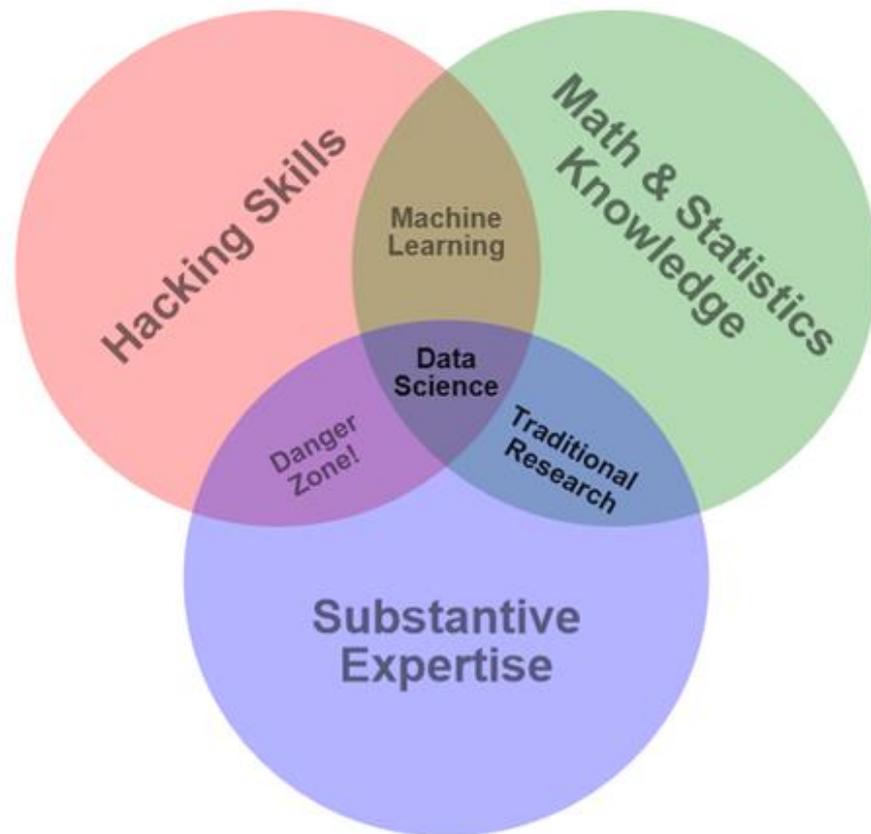
All models are wrong, but some are useful.

George E. P. Box

# 模型的组成有哪些？



Data Mining



Data Science

回顾过去**数据挖掘**的发展到**大数据**，整体而言，统计学科的本质是没有变的，**分析的核心观念**没有因为数据量的多寡而有改变，而是应用方面更强调交叉学科(学科间的协作)

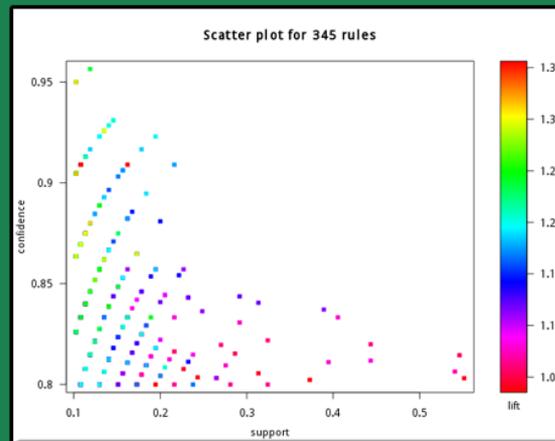
# 分析的组成要素



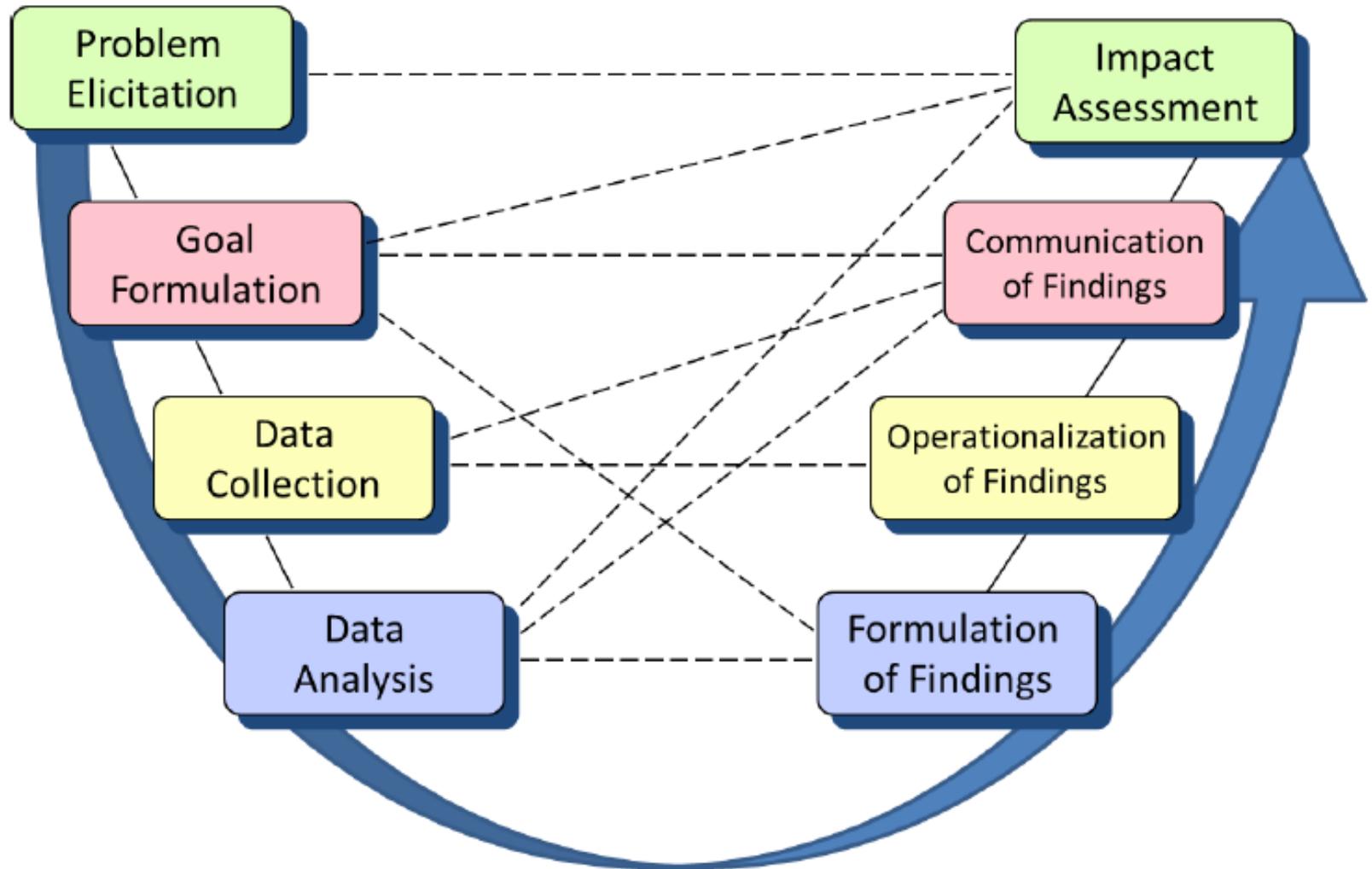
数据分析与导引系统  
Data Analysis & Guiding System

首页 | 网站导航 | 会员登录 | 新手使用者

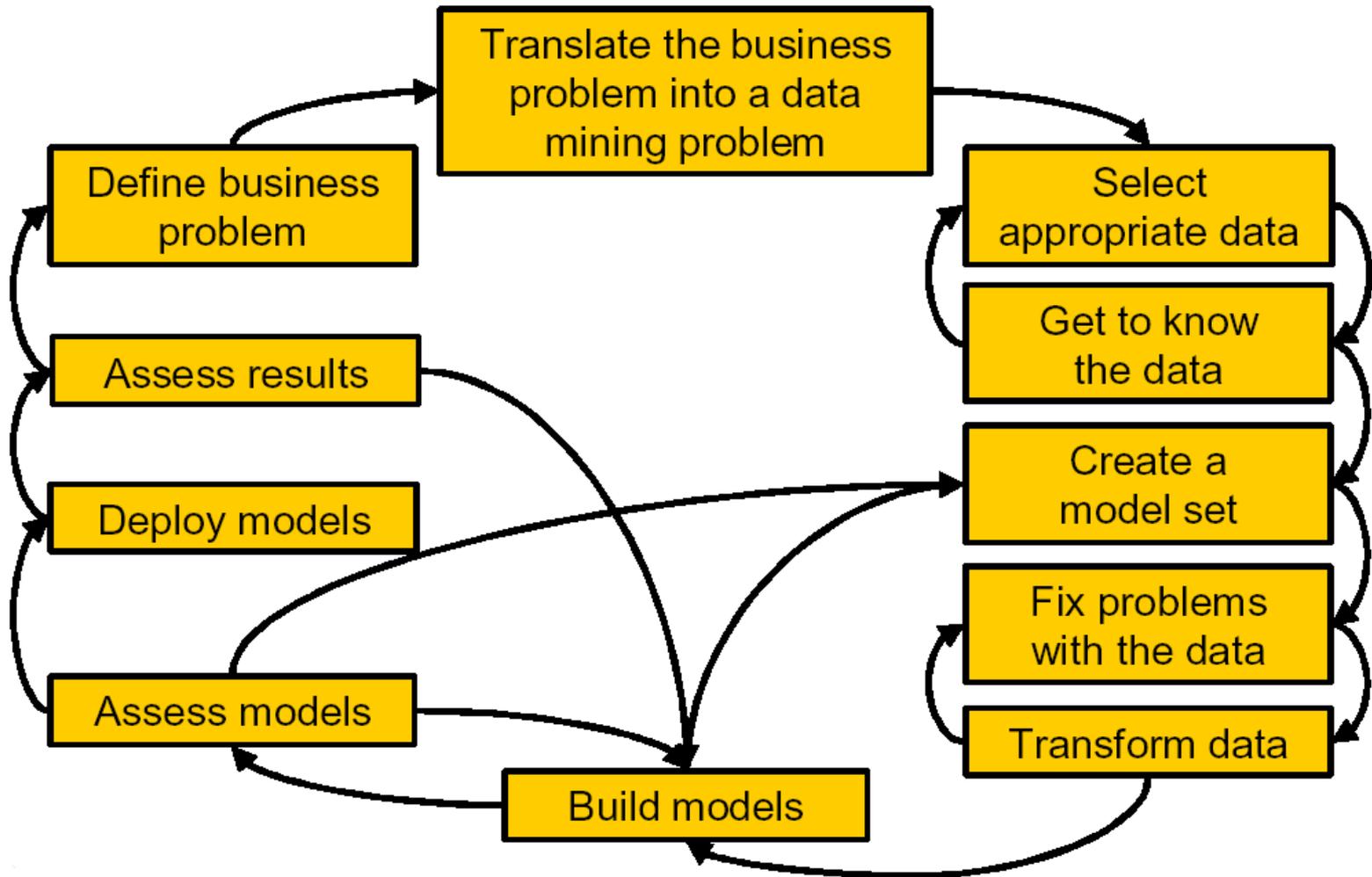
新手导引 数据处理 分析方法 图表绘制 输出结果



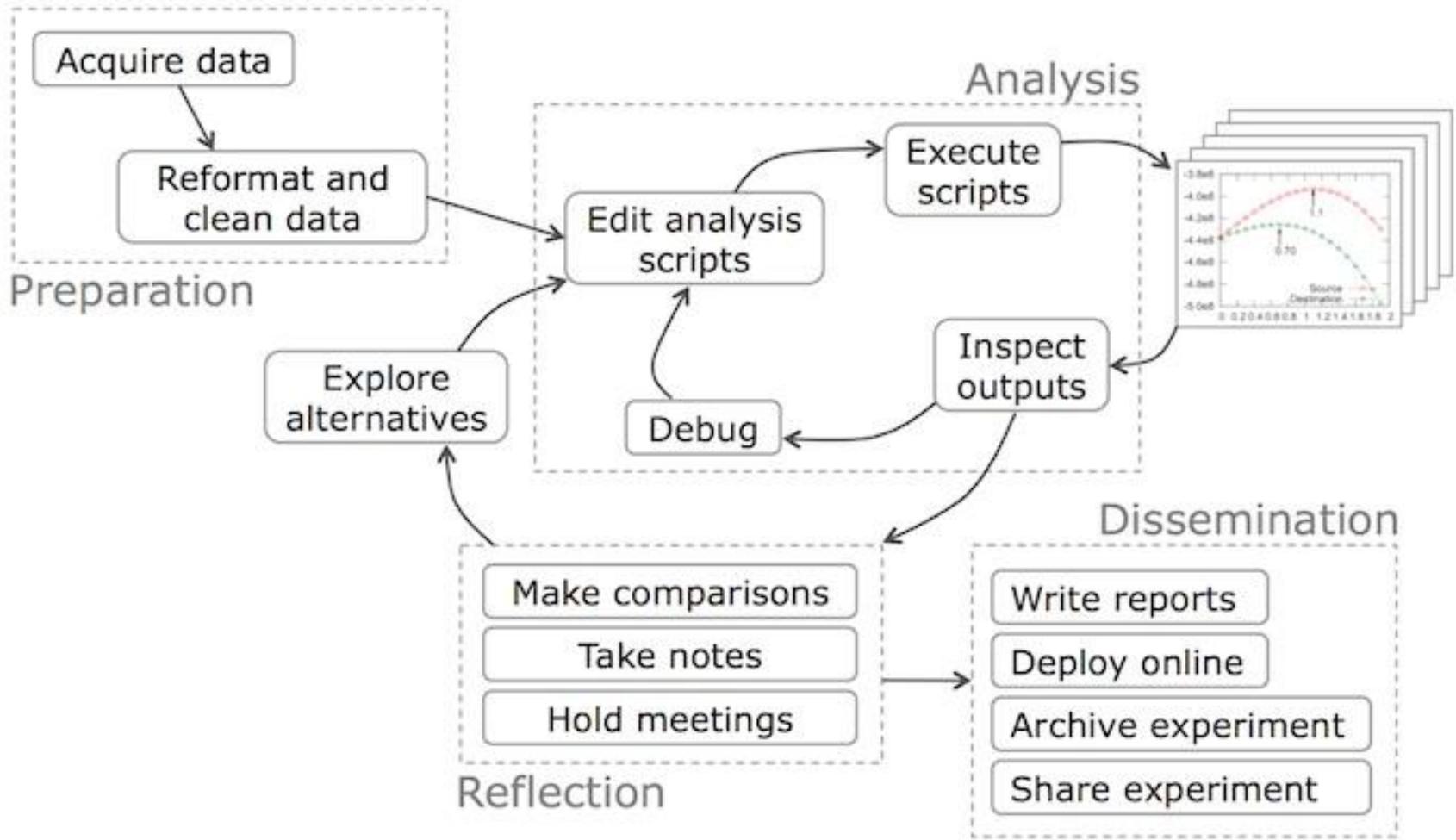
# Statistics: A Life Cycle View



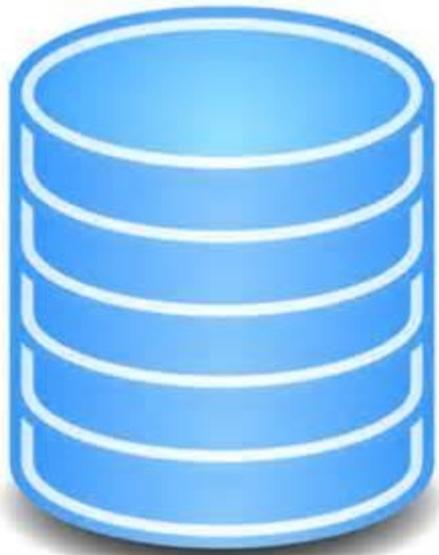
# Data Mining Is Not a Linear Process



# The Data Science Workflow



回想你做(大)数据分析的过程,  
那些是重要的要素?



数据

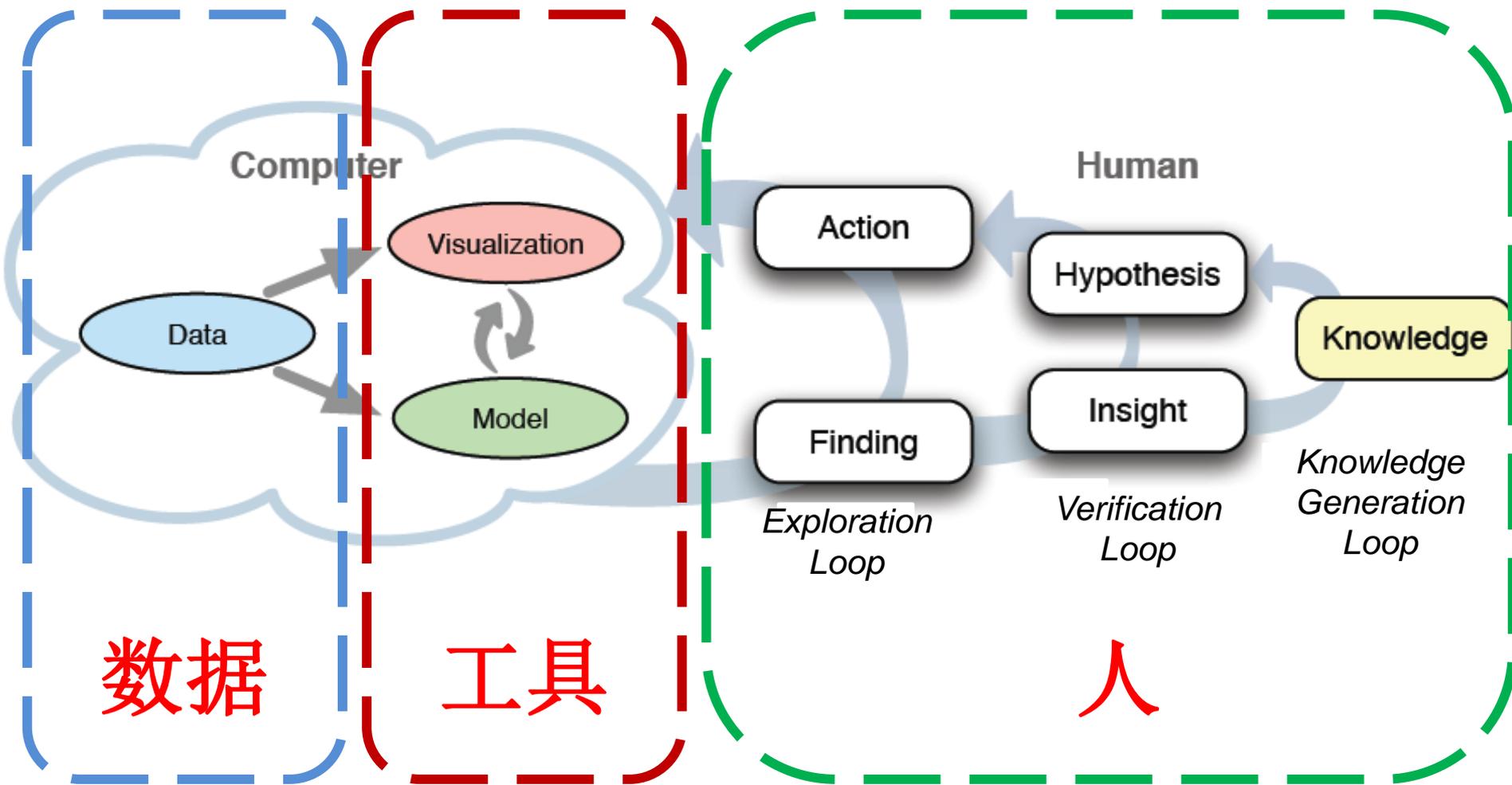


工具



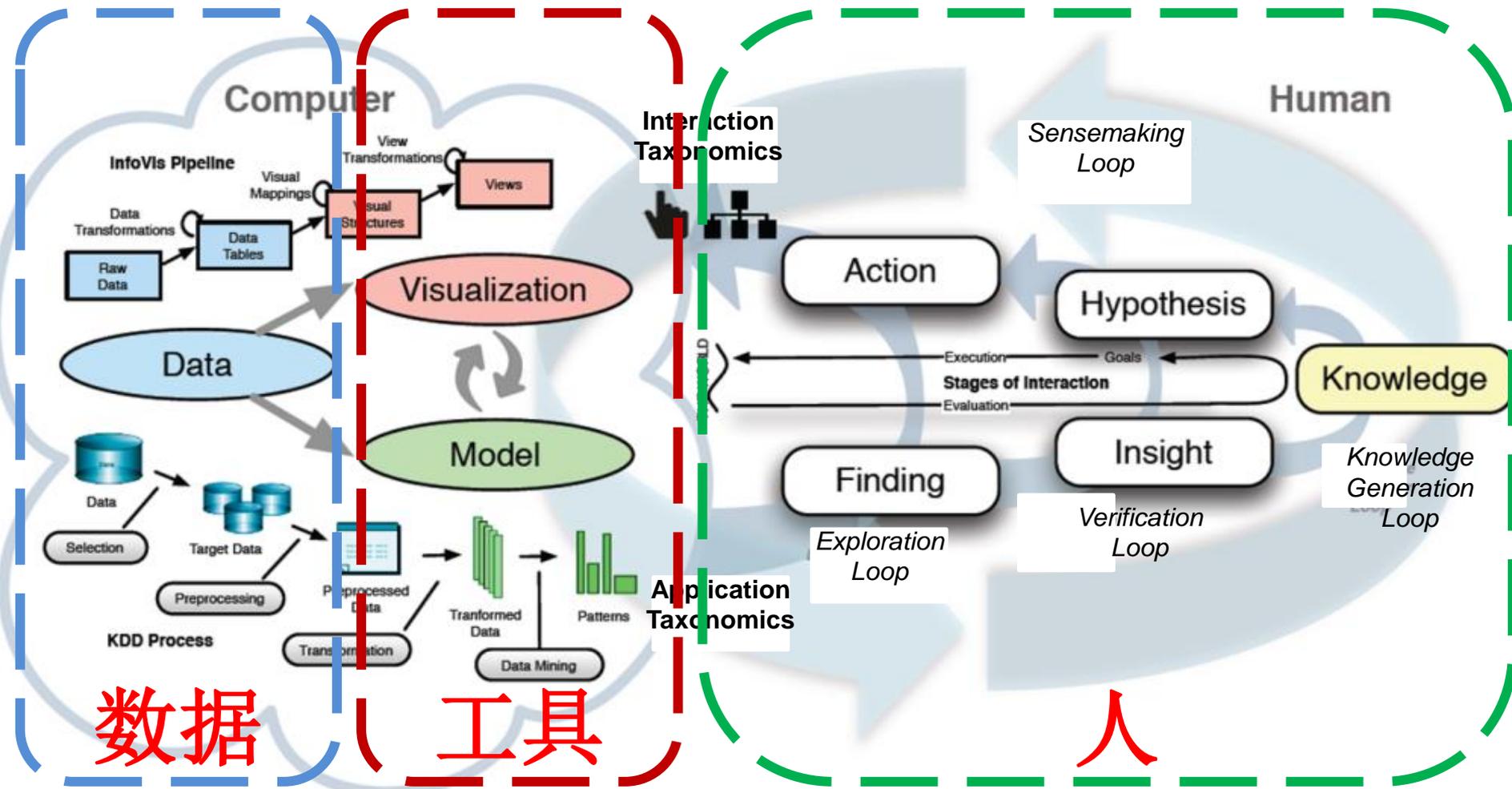
人

# 数据分析中的知识产生过程(模型)



Source : Knowledge Generation Model for Visual Analytics, IEEE Trans. Vis. Comput. Graph. (VAST'14)

# 数据分析中的知识产生过程(模型)



Source : Knowledge Generation Model for Visual Analytics, IEEE Trans. Vis. Comput. Graph. (VAST'14)

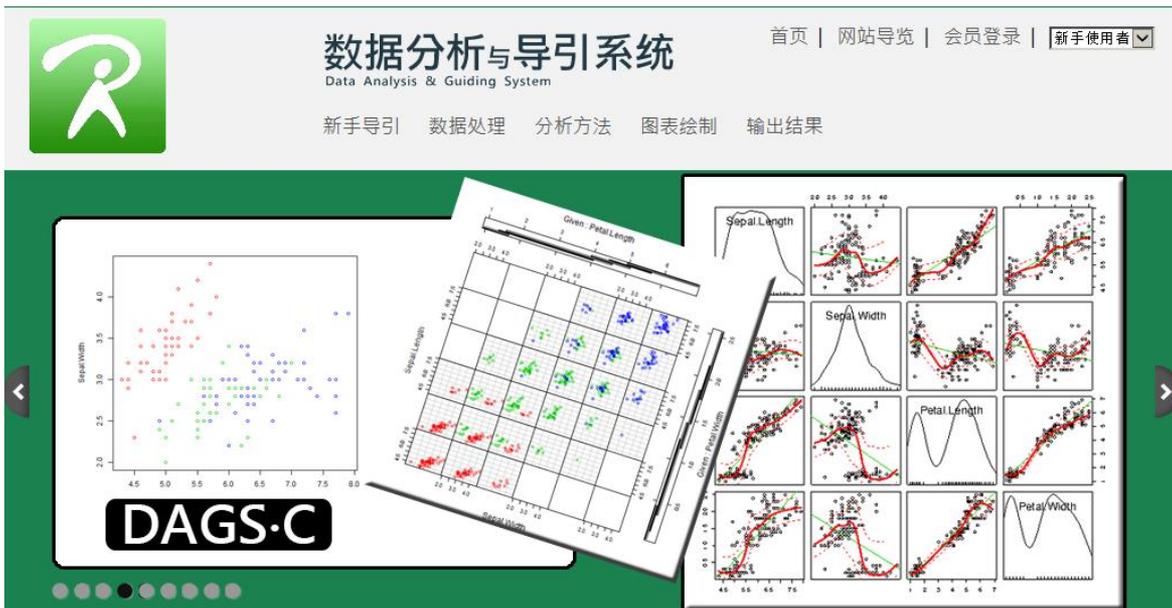
# 平台设计理念

r-web.sysu.edu.cn

The screenshot displays the user interface of the 'Data Analysis & Guiding System' (DAGS-C). At the top, there is a navigation bar with a logo on the left and links for '首页', '网站导航', '会员登录', and '初阶使用者'. Below the navigation bar, a menu contains '初阶导引', '数据处理', '分析方法', '图表绘制', '概率分布', and '输出结果'. The main content area is divided into two panels. The left panel, titled '时间数列(Time Series)', features a line graph of a time series and the text '分析过去时间数列数据 帮助预测未来变化趋势'. The right panel, titled 'Series income (nlag = 24)', shows an Autocorrelation Function (ACF) plot with a y-axis labeled 'ACF' and an x-axis labeled 'Lag'. The ACF plot includes a blue dashed horizontal line at approximately 0.15 and a vertical dashed line at lag 24. An inset window in the top right of the ACF plot shows a zoomed-in view of the time series data.

# 平台设计理念

1. 降低分析工具的使用门坎
2. 让数据分析方法能为大家所熟悉及使用
3. 增强企业分析效率
4. 方便人员协作互动



# 平台系统演示

r-web.sysu.edu.cn



数据分析与导引系统  
Data Analysis & Guiding System

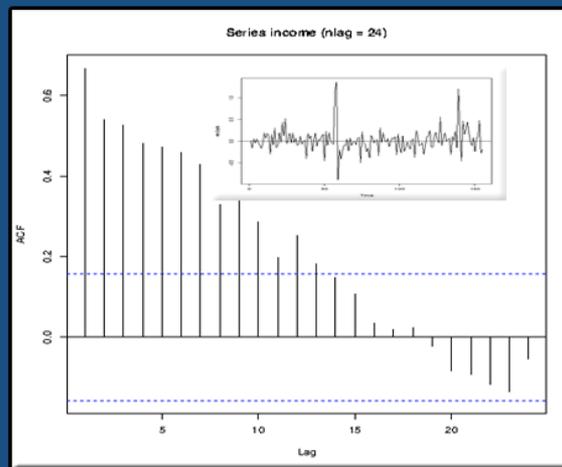
首页 | 网站导航 | 会员登录 | 初阶使用者

初阶导引 数据处理 分析方法 图表绘制 概率分布 输出结果

## 时间数列(Time Series)

分析过去时间数列数据  
帮助预测未来变化趋势

DAGS·C



# 未来工具的设计也需要与时俱进的改进

## 平台在探索循环的优点

1. 引导性强
2. 范例简单易懂
3. 引导用户快速系统的发现数据中的模式
4. 清楚的视频教学指引

## 平台在验证循环的优点

1. 研究者可以如何保存之间的探索结果，以方便回溯，验证其可靠性
2. 研究者可以考虑如何组织不同的探索结果，辅助用户产生假设，甚至自动产生假设

# 未来工具的设计也需要与时俱进的改进

## 未來的研發路徑

1. 通过可视化与数学模型进行交互
2. 如何自动检测模式
3. 考虑不同算法之间的协作
4. 考虑分析人员和领域专家之间的沟通
5. 考虑动态变化的流数据
6. R代码的生成
7. 加强数据可视化
8. 不同算法的组合
9. 使用者自定义方法并集成个性化的解决方案

这个平台正在演进，欢迎学校和企业探讨不同的合作模式！！

林禎舜 (Eric Lin)

Mobile : 13911820560

E-mail : [linchenshun@gmail.com](mailto:linchenshun@gmail.com)