

Big Data

School of Applied Science and Civil Engineering
Department of Data Science

H.B.Sun

2018.01.17

Big Data

- Warm up
- Brief Introduction
- What and How Should We Do?

Warm Up

Video-a bit long

Brief Introduction

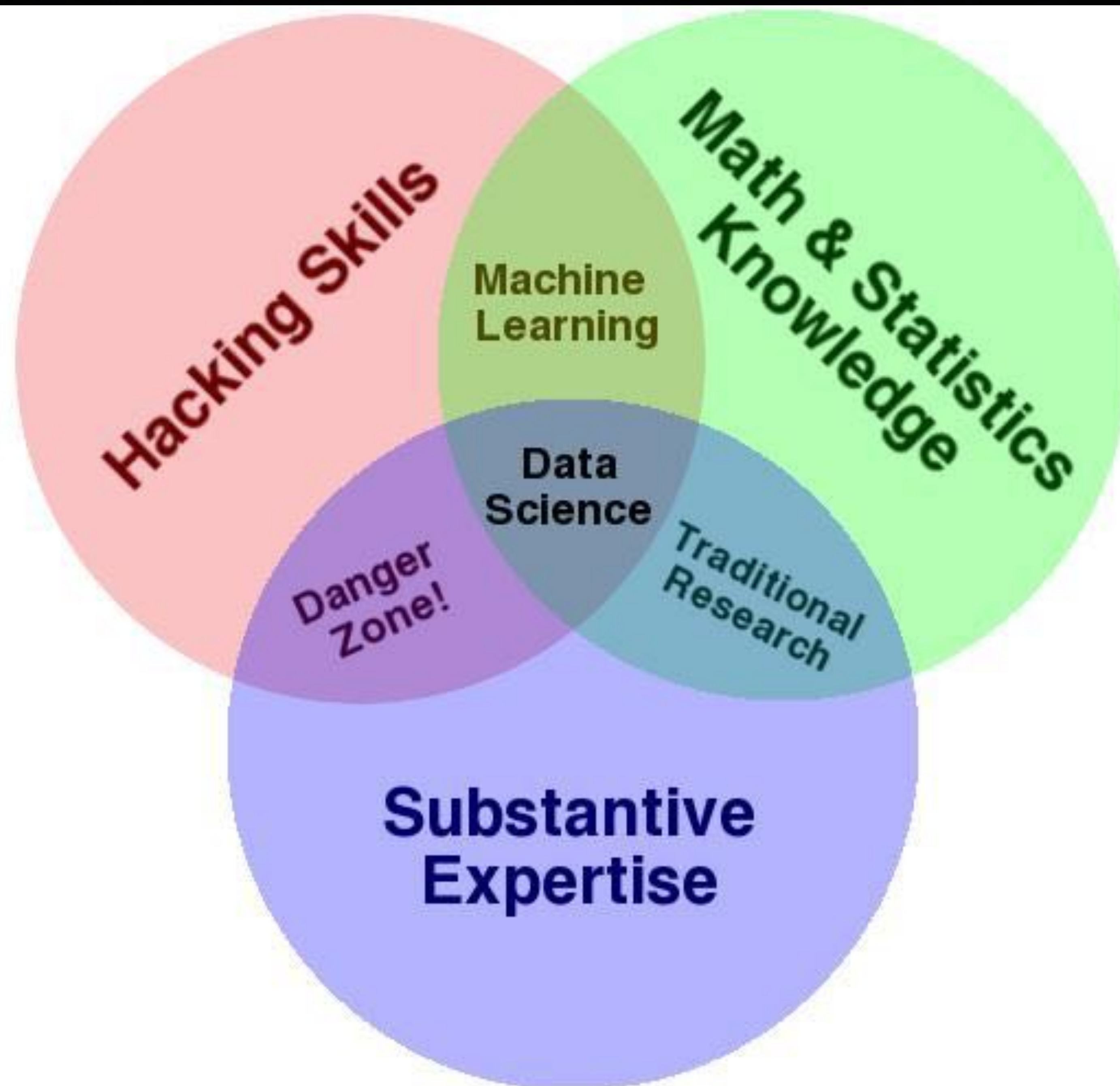
- Big Data isNot only big!
- 对数据定义的延伸迫使统计学与计算机科学从各自的传统核心领域进入边缘领域，两个学科在边缘领域的融合成为数据科学的基础。

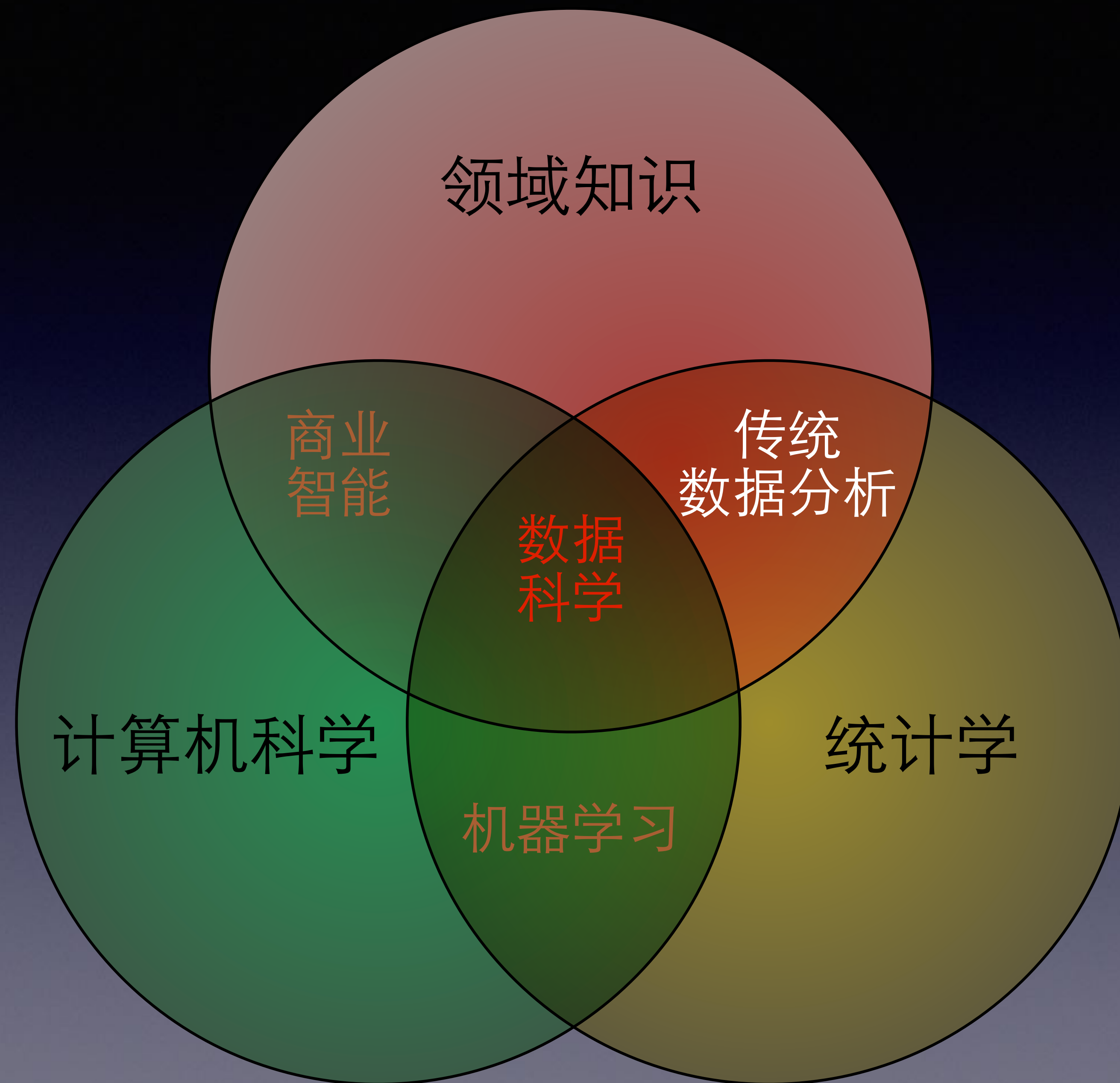
Brief Introduction

- Data Scientist is?
- 擅长统计学的软件工程师与擅长软件工程的统计学家的混合体；
- 擅长对数据分析结果进行解释的科学家与擅长数据分析的 MBA 混合体；

Brief Introduction

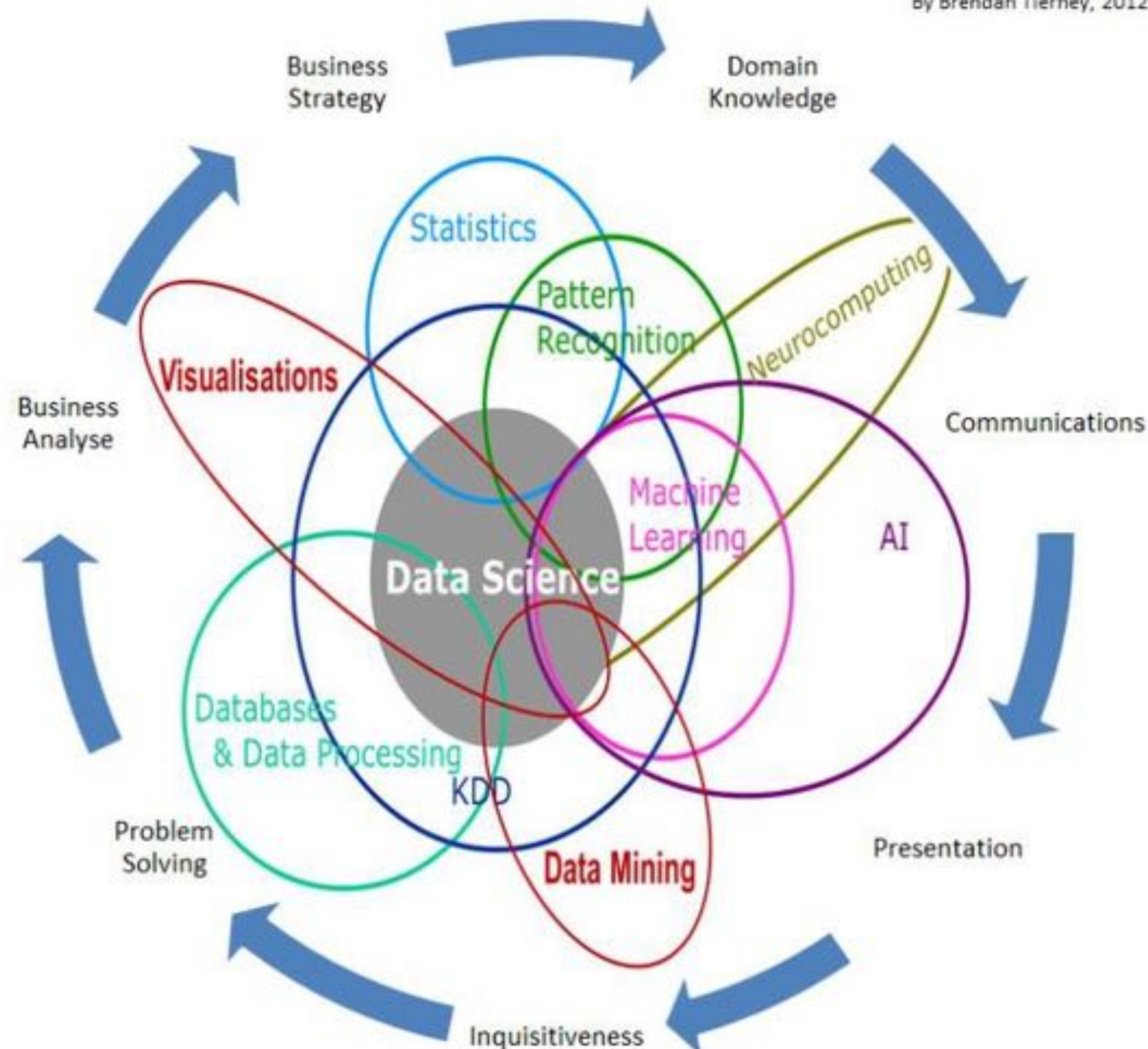
- Statistics and Data Science.....?
- 统计学的大多数传统核心内容还未做好处理以十亿为单位的记录及非结构化数据的准备。同样，计算机科学的核心领域虽然擅长高效获取并存储大量结构化和非结构化数据，但在通过建模、分类和可视化等形式将数据转换成信息的能力上存在短板。
- 虽然数据科学不是统计学，但统计学为这门学科贡献了基础方法。



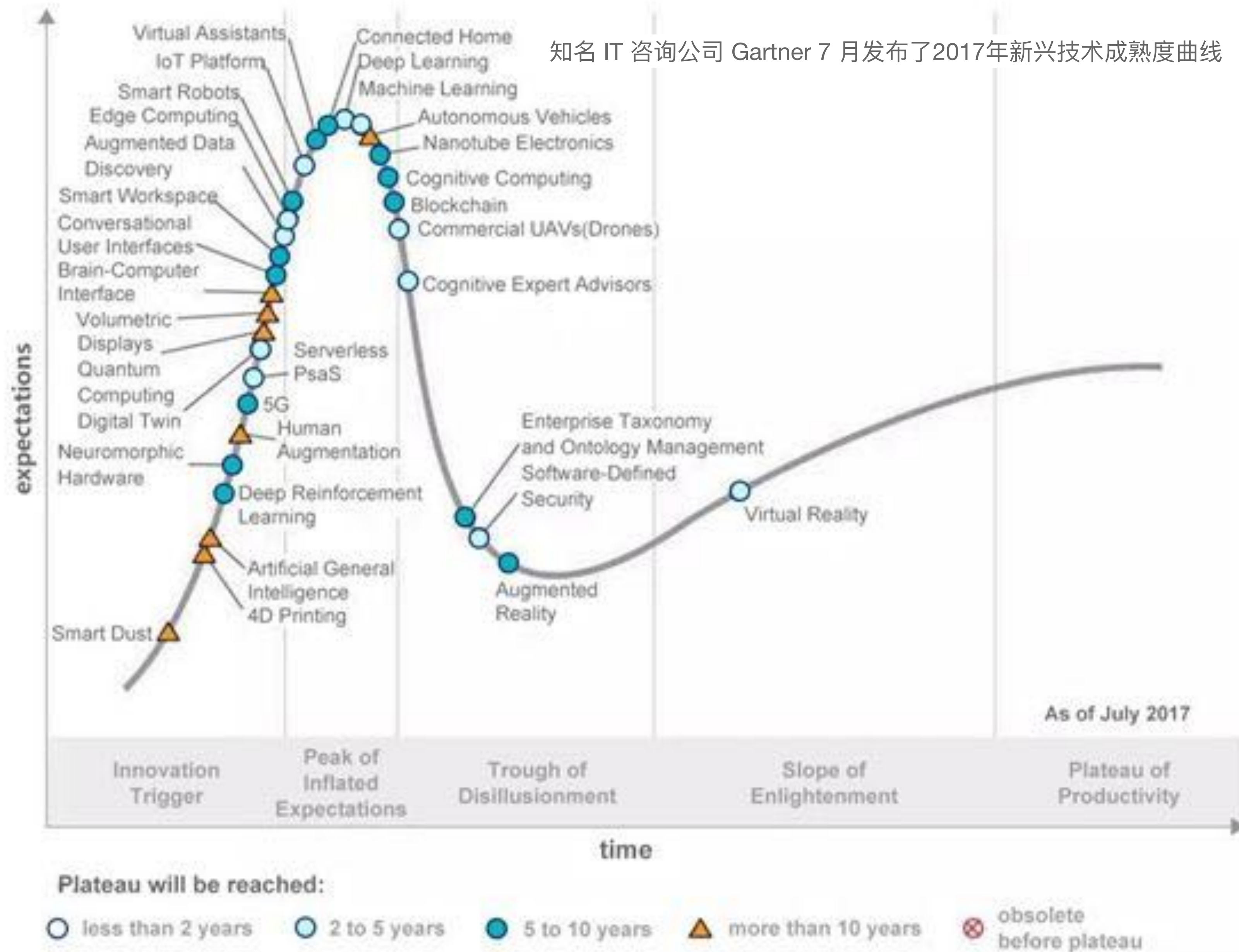


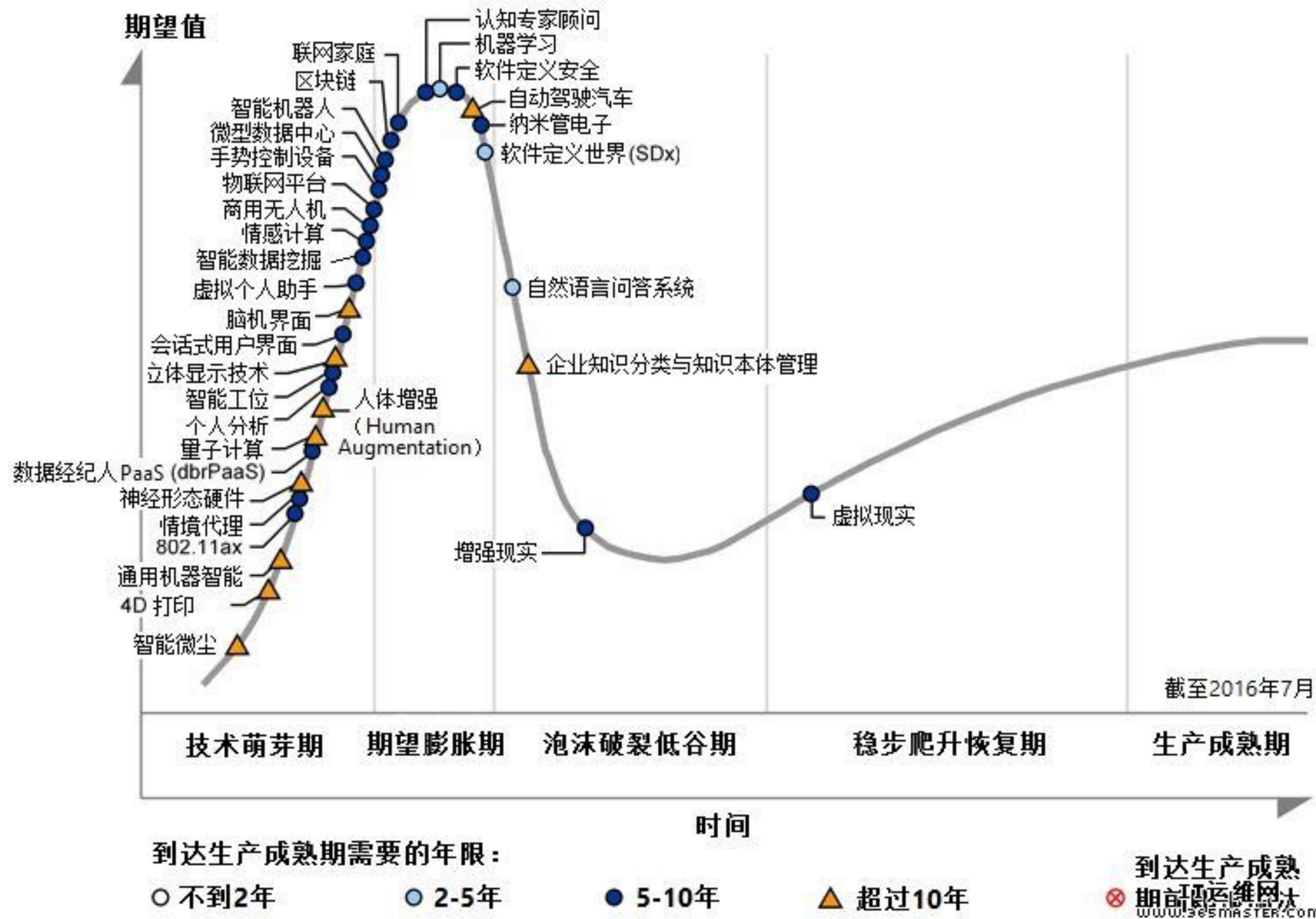
Data Science Is Multidisciplinary

By Brendan Tierney, 2012



知名 IT 咨询公司 Gartner 7 月发布了2017年新兴技术成熟度曲线



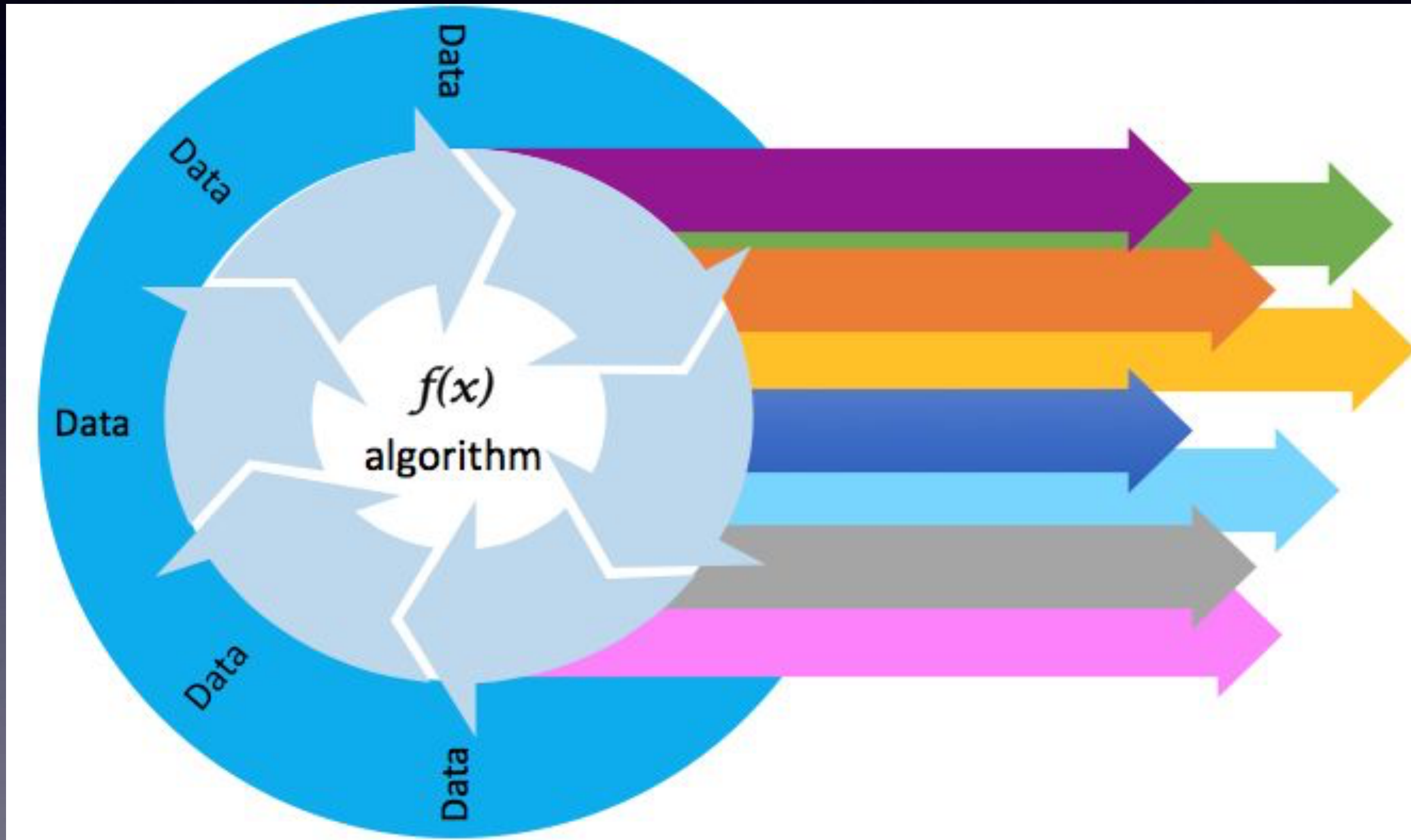


Big Data Infrastructure

Big Data Applications



What and How Should We Do?



What and How Should We Do?



O'REILLY®



图灵程序设计丛书



数据科学入门

Data Science from Scratch

[美] Joel Grus 著
高蓉 韩波 译



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS

O'REILLY®



Data Science from Scratch

FIRST PRINCIPLES WITH PYTHON

Joel Grus

www.allitebooks.com

Thank you for your patient!

to be continued