

接近理科博士生水准！OpenAI放大招

来源：证券时报网 作者：周春媚 2024-09-13 10:09

北京时间9月13日凌晨，OpenAI重磅发布全新AI大模型——o1模型。

据OpenAI官网介绍，这一模型“旨在花更多时间思考后再作出反应。它们可以推理复杂的任务，解决比以前的科学、编码和数学模型更难的问题”。不过，今天在ChatGPT和大模型API中新发布的是该系列中的第一款模型，而且还只是预览版——o1-preview（o1预览版）。

“在我们的测试中，下一个模型（o1模型）更新在物理、化学和生物学的具有挑战性的基准任务上的表现与博士生相似。我们还发现它在数学和编码方面表现出色。在国际数学奥林匹克(IMO)资格考试中，GPT-4o仅正确解决了13%的问题，而推理模型得分为83%。他们的编码能力在比赛中得到了评估，并在Codeforces比赛中达到了第89个百分位。”OpenAI对o1模型的介绍表示。

在数学竞赛中，以AIME2024为例，GPT-4o平均只能解决12%的问题，而o1平均能解决74%的问题，若采用64个样本的共识，解决率能达到83%。o1在竞争性编程问题（Codeforces）中排名第89位，在美国数学奥林匹克(AIME)预选赛中跻身美国前500名学生之列，并在物理、生物和化学问题(GPQA)基准测试中超越人类博士级准确度。



*"It is not the strongest of the species that survive,
nor the most intelligent,
but the one that is most **responsive to change**."*

- Charles Darwin

“能够生存下来的物种，不是那些最强大，也不是那些最聪明，而是那些对变化反应最快的物种。”

- 达尔文

郑磊(George) – 战略与创新创业系 副教授

- 2005 ~ 2010 新加坡国立大学 学士 应用数学(计算机)
- 2015 ~ 2020 新加坡管理大学 博士 战略与组织管理
- 五年咨询业从业

研究方向：

- 战略管理
- 企业行为理论(卡内基学派): 组织设计、复杂系统

兴趣爱好：网球、带娃

Email: zhenglei@shufe.edu.cn

Office: 商学院508

Office Hour: 周一全天、二下午、周三上午

助教: 待定

卡内基学派: 赫伯特-西蒙, 詹姆斯-马奇



卡内基学派

司马贺 (赫伯特西蒙 Herbert Simon)

- 1975年图灵奖、1978年诺贝尔经济学奖
- 首批中科院外籍院士 (1994)



管理学 - 组织理论

理查德赛特 (Richard Cyert)

- 卡内基大学校长

马奇 (James March)

- 加州大学社科学院首任院长

经济学 - 有限理性

威廉姆森 (Oliver Williamson)

- 2009年诺贝尔经济学奖

戴尔·莫滕森 (Dale Mortensen)

- 2010年诺贝尔经济学奖

计算机 - 人工智能

费根鲍姆 (Edward Feigenbaum)

- 1994年图灵奖得主

在有限理性下，组织需要如何更好地做决策？

- 此时没有最优解
- 多人决策优化
- 设计“聪明的组织”
- 应对变化的环境

简短的自我介绍：

- 姓名
- 专业/年级

分享：

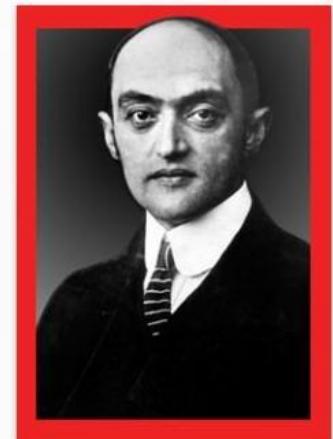
- 创新创业的点子/创新创业的经历/比赛
- 或者，一个发现身边的改变生活的创新

Entrepreneurs will seek to use technological innovation, a new product/service or a new process for making it, to get strategic advantage.

- Monopoly profits 技术垄断的租金
- Imitation 模仿与反模仿
- Equilibrium 打破平衡
- Creative destruction 创新性破坏

Joseph Schumpeter

熊彼特



创新的形式

1. 新产品



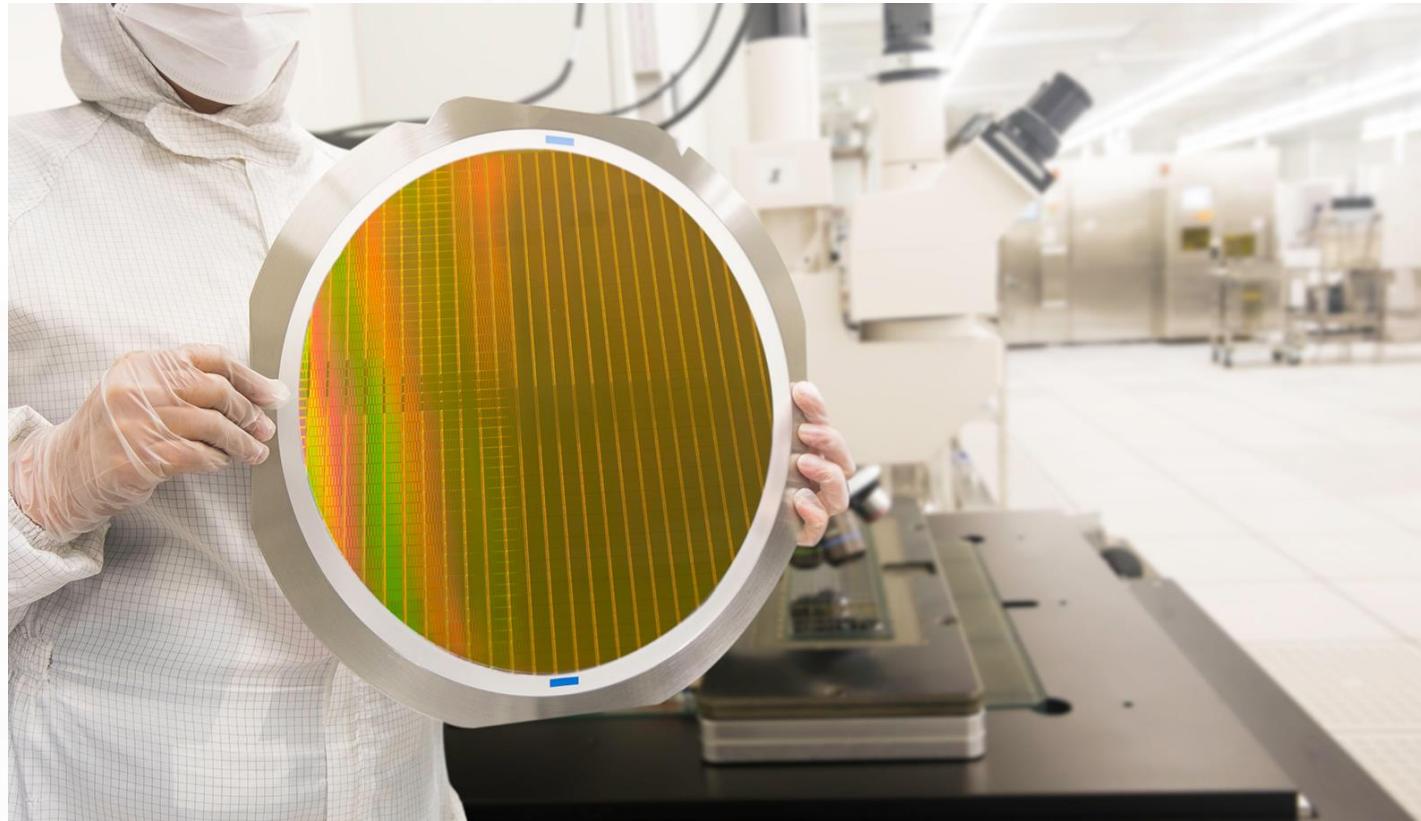
创新的形式

2. 新流程



创新的形式

3. 复杂性



创新的形式

4. 法律法规



创新的形式

5. (恰当的) 时机



创新的形式

6. 平台/系统



创新的形式

7. 新的协议



IPv4

vs.

IPv6



创新的形式

8. 从别的领域借鉴



Tennis Strings



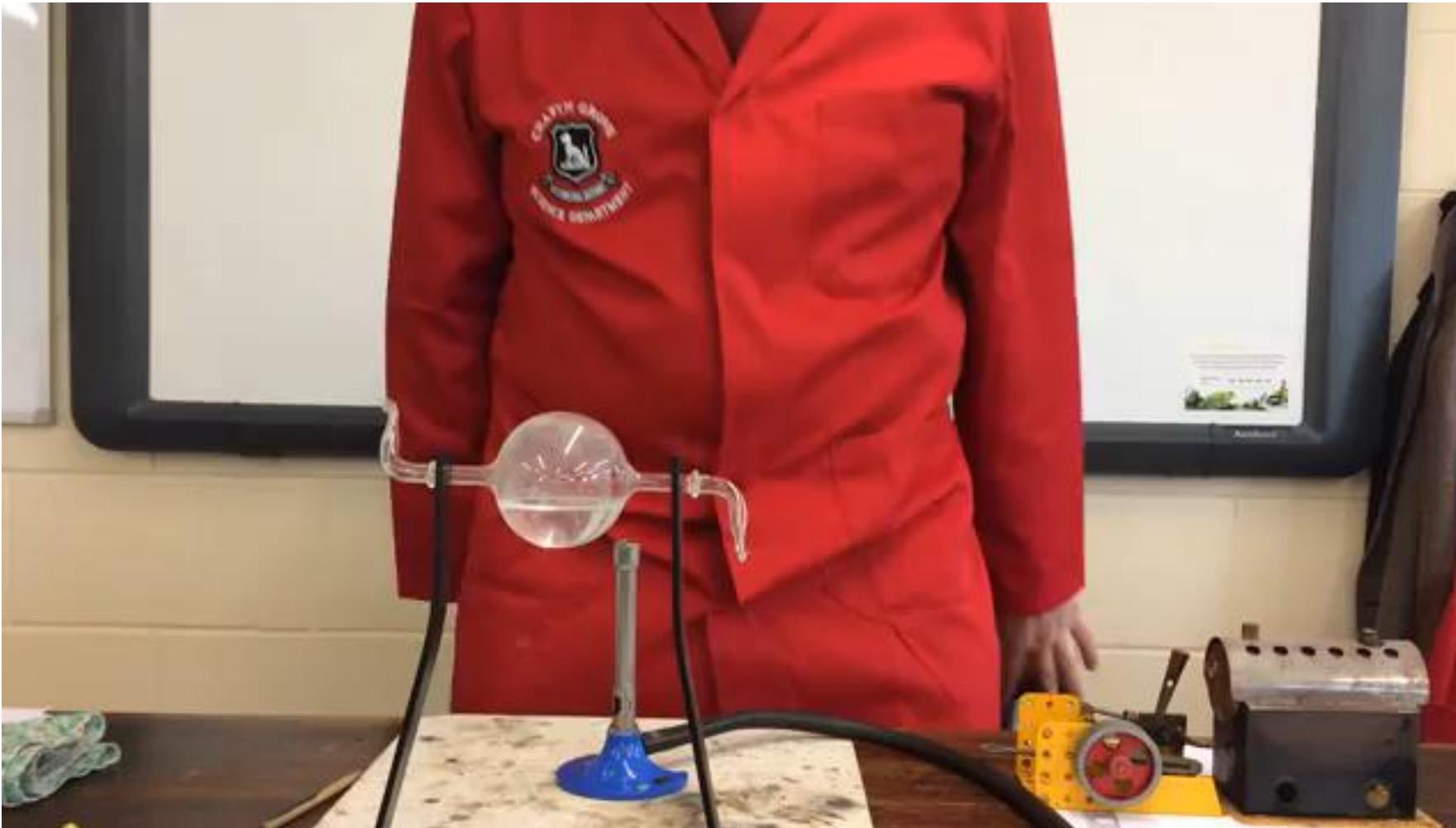


工业革命：科学、技术、科技创新

Hero's Engine

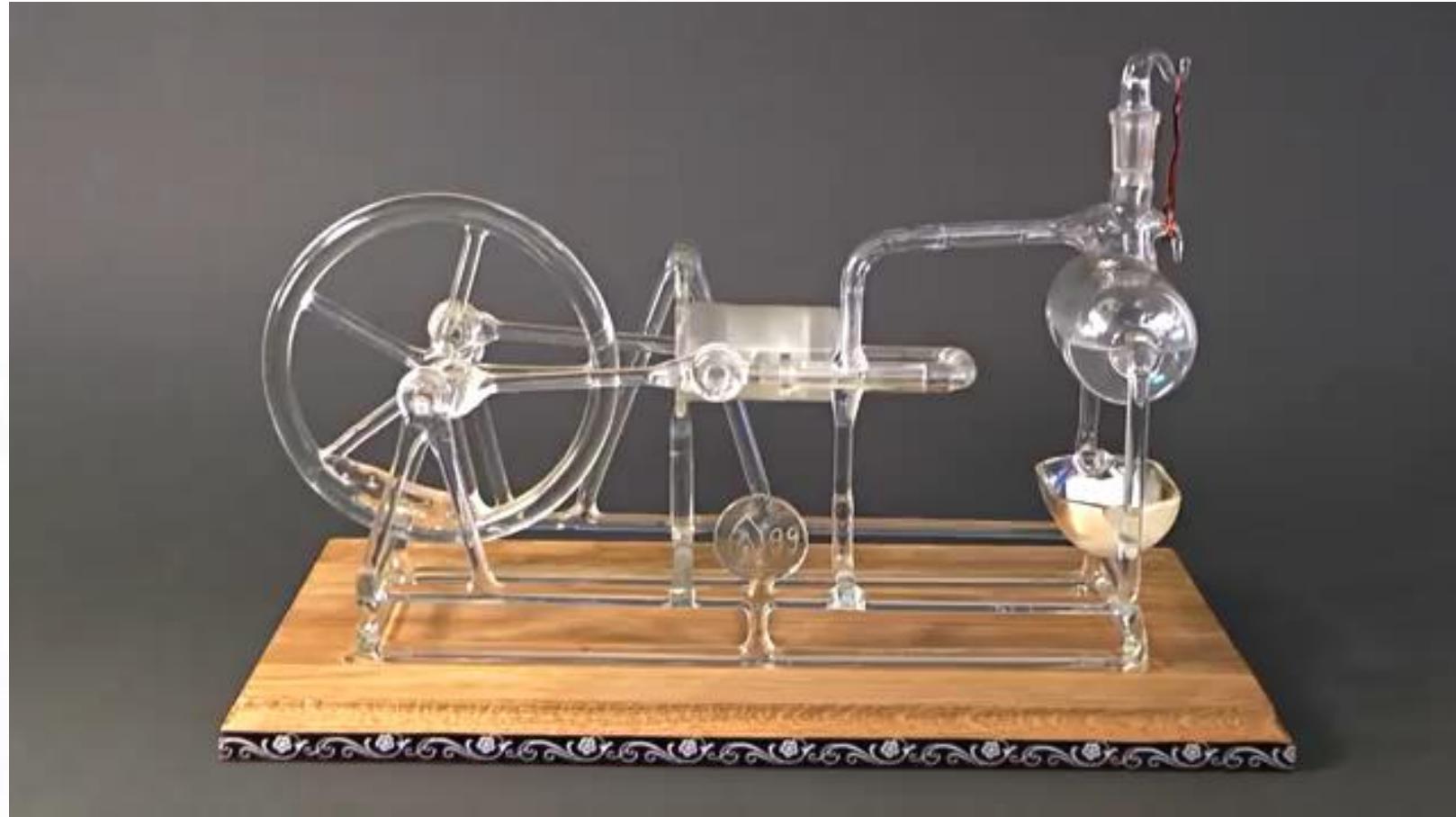
MIT Department of Physics
Technical Services Group

工业革命：科学、技术、科技创新



工业革命：科学、技术、科技创新

史蒂芬森蒸汽机



【全网最清晰的汽车发动机工作原理4K高清3D动画】

<https://www.bilibili.com/video/BV1er4y1J7dH/>

月球知识局 bilibili

原理知识局





科学、技术、科技创新

科学Science:

- 基于或以科学方法和原理为特征，指通过观察和实验系统地研究物理和自然世界的结构和行为的智力和实践活动

技术Technology:

- 将科学知识应用于实际目的，尤其是在工业领域
- 应用科学知识开发的机械和设备
- 涉及工程或应用科学的知识分支

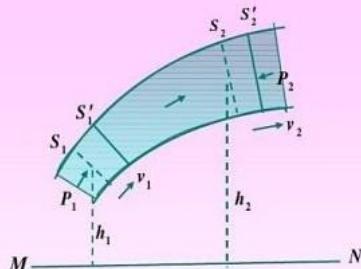
科技创新：

- 创造新的科学技术
- 在生活生产中运用新的科学技术

科学↔技术



- ❖ 伯努利方程
- ❖ 理想流体在同一管道中稳定流动时，流体压强、流速和高度所具有的压强能、动能及位置势能是恒定的。即，伯努利方程是流体力学中的能量转换与守恒定律。



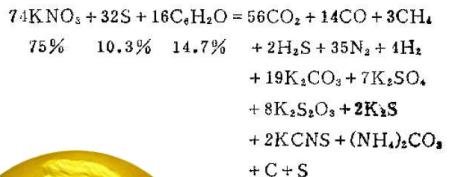
$$p_1 + \rho \frac{v_1^2}{2} + \rho g h_1 = p_2 + \rho \frac{v_2^2}{2} + \rho g h_2 = \text{常数}$$

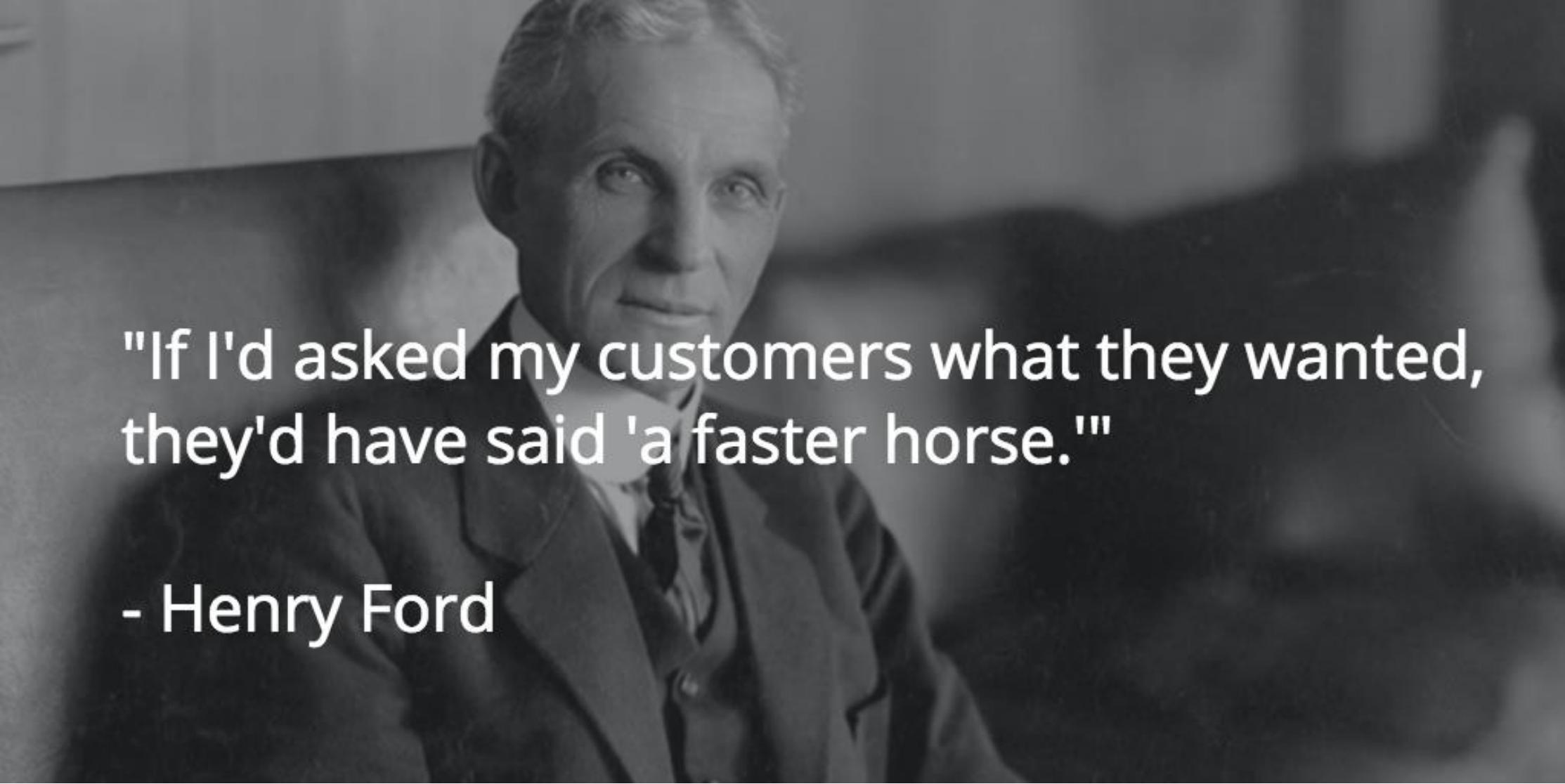
应用伯努利方程，人们制成了文丘里流量计及水流抽气等应用。

船舶 航空航天 车辆工程 电子电气 机械设计



黑火药燃烧时的化学反应是极为复杂的，所以反应方程式也十分复杂，Casete 提出的黑火药燃烧反应方程式为：

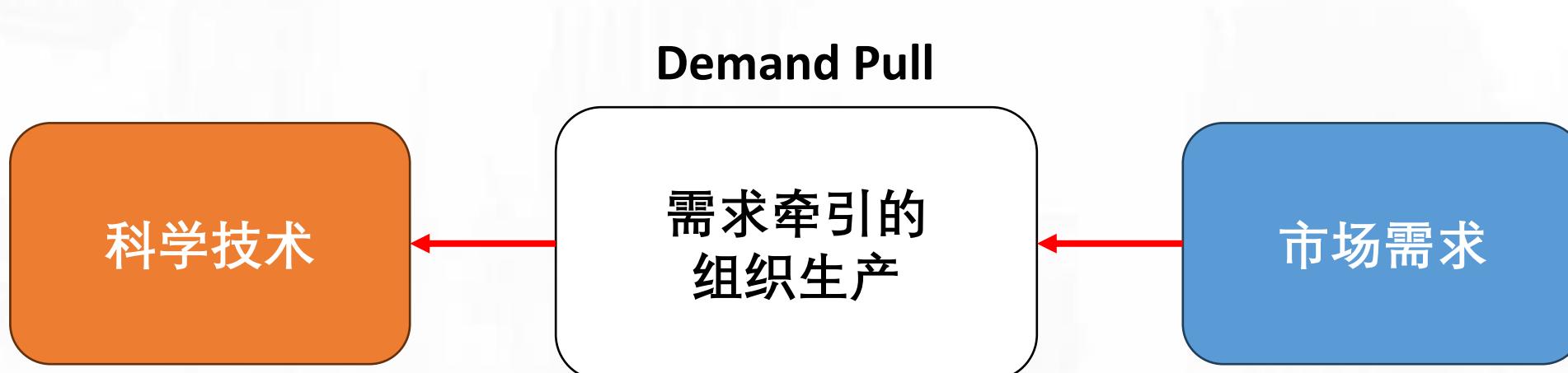
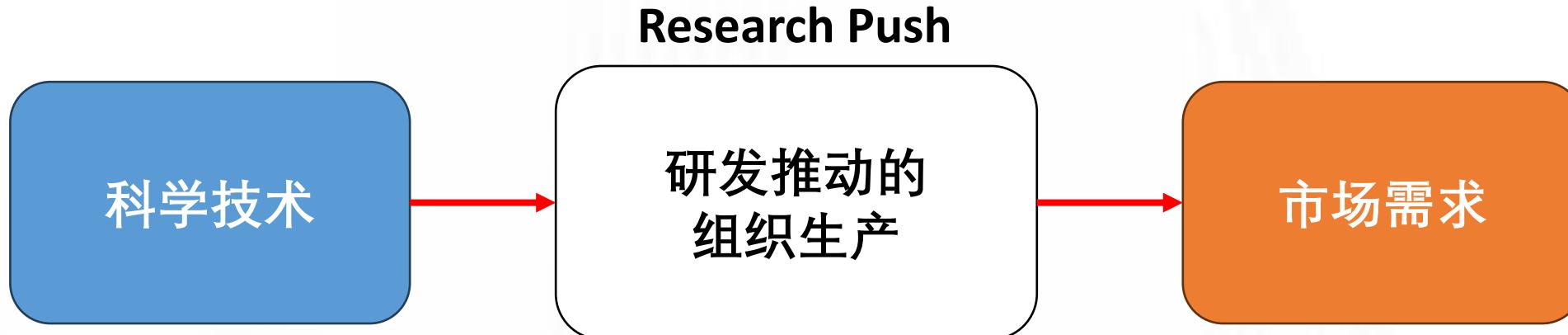




"If I'd asked my customers what they wanted,
they'd have said 'a faster horse.'"

- Henry Ford

Research Push VS Demand Pull





技术↔创新 (课程内容)

- Artificial Intelligence (AI)
 - AIGC
 - LLM / GPT
- Virtual Reality (VR)
 - Glasses and 3D Modelling
- 3D printing
- Livestream



创新↔实践（课程内容）

- 企业/校外参访（待定）

小游戏：产品设计

屏幕 0-瀑布 1-刘海	尺寸 0-八寸 1-五寸	解锁 0-指纹 1-面容	绩效
0	0	0	0
1	0	0	-1
0	0	0	-2
1	1	1	-2
0	1	1	-2
1	0	1	-2
1	1	1	3

手机是一个复杂的产品，单个部件（屏幕、尺寸、解锁方式、颜色等）对手机性能以及消费者满意度的影响收到其他部件的影响！



回顾：复杂性

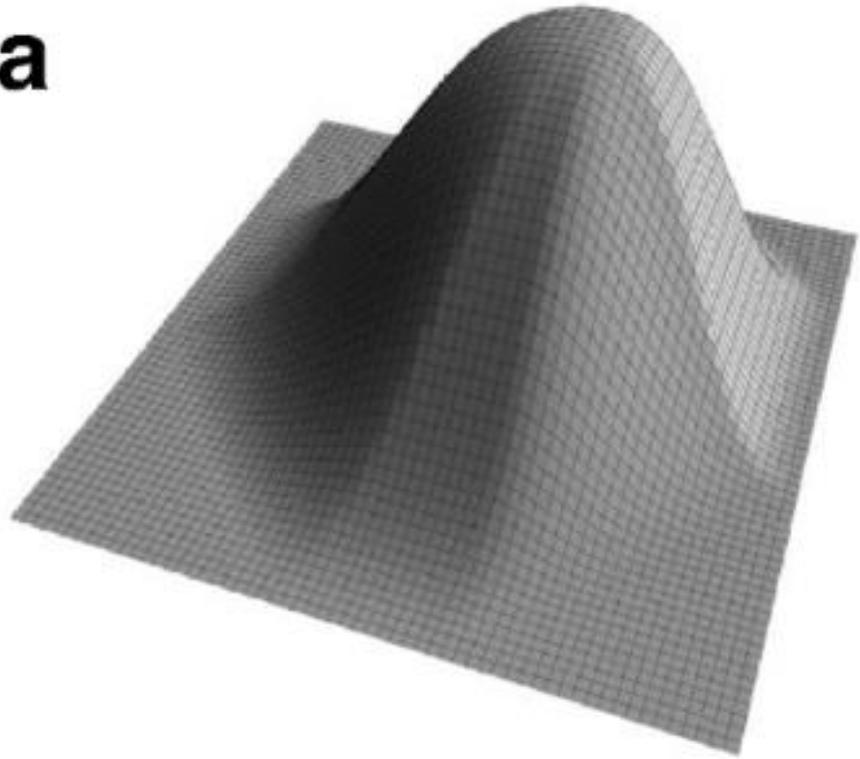
当一个系统是由**多个部件**组成，并且：

1. 局部部件对全局系统的影响是**非线性的(Nonlinear)**——既部件对全局的影响无法被展示成部件输入的线性方程，因此，
2. 系统具有**突发性(Emergent)**——全局系统的行为无法直接通过局部的部件输入进行预测

这样的系统具有**复杂性(Complexity)**

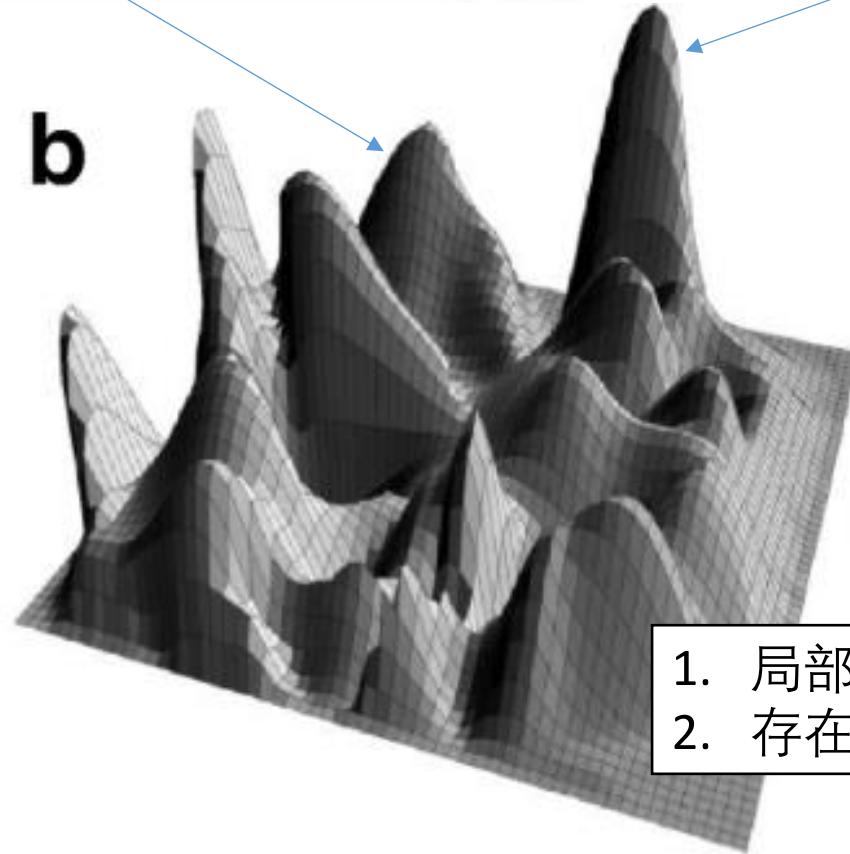
复杂性与绩效平面

a

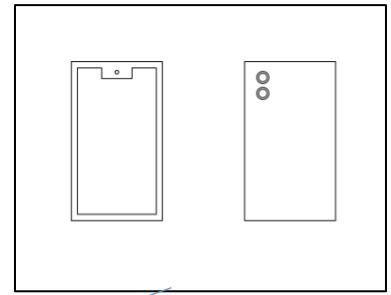
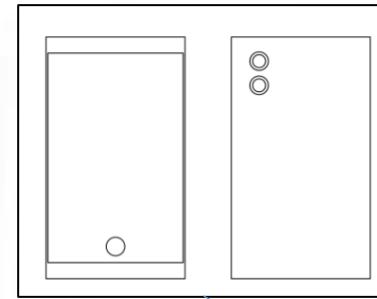


低复杂性的绩效平面

b



高复杂性的绩效平面



- 1. 局部最优与全局最优
- 2. 存在多个局部最优解