



认知图谱

—人工智能的下一个瑰宝

唐杰
清华大学

人工智能的发展

符号AI

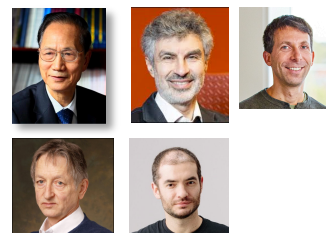
1956 Dartmouth Conference:
The Founding Fathers of AI



符号模型/
规则模型/
感知机

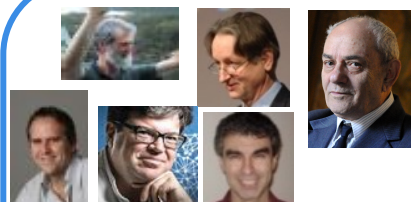
第一代

认知智能



张钹院士2016年提出第三代人工智能雏形，DARPA 2018年发布AI Next计划。核心思路是推进数据统计与知识推理融合的计算；与脑认知机理融合的计算。

第三代

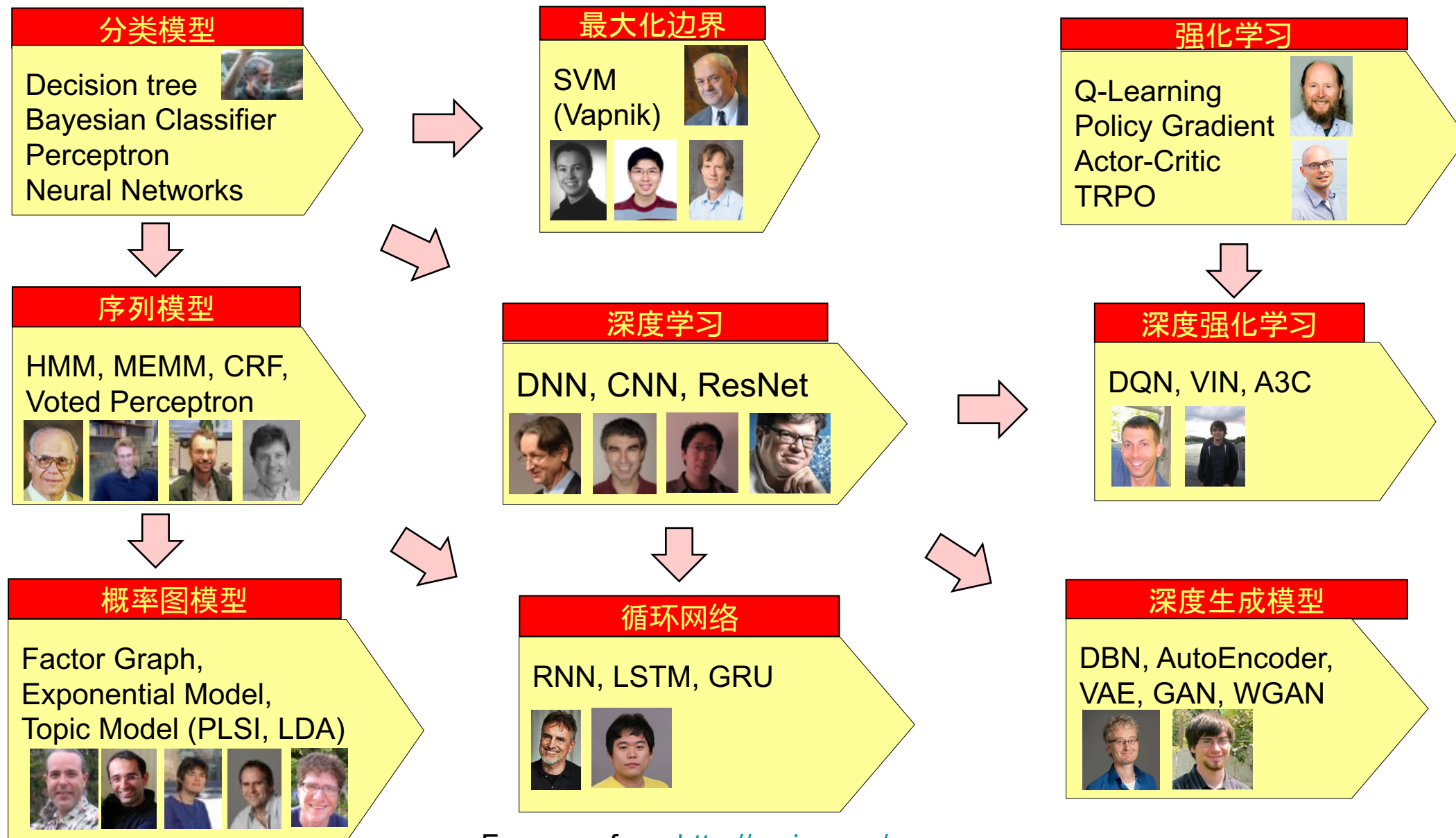


大数据驱动的统计学习方法初步实现了针对文本、图像、语音等的感知与识别

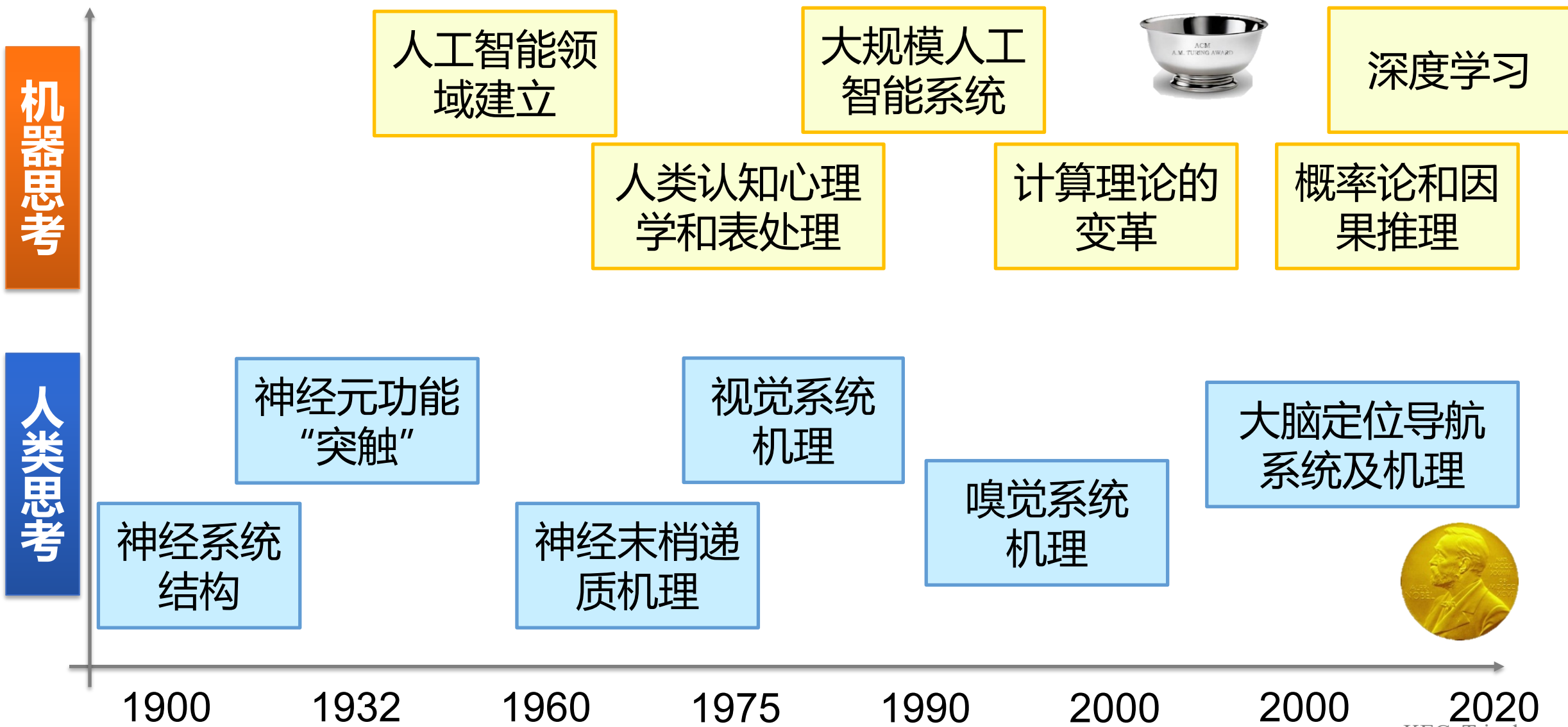
感知智能

目前急需的是高质量超大规模知识图谱（**AI的基础设施**）以及对超大规模数据的**深度理解能力**（面向认知的深度学习）

回顾机器学习

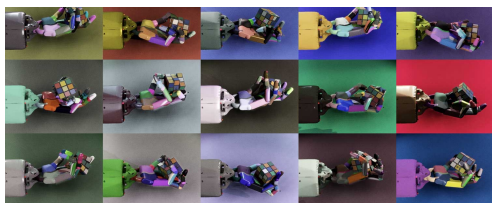


机器思考 vs. 人类思考

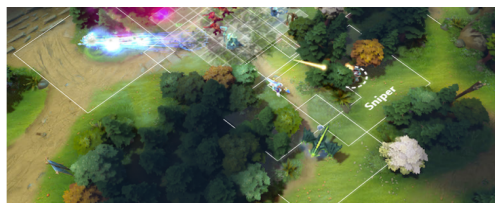


迈向通用人工智能

- OPEN AI的通用人工智能战略：制造通用机器人和使用自然语言的聊天机器人



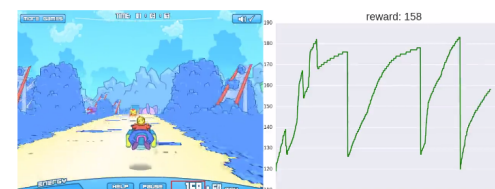
单个机械手还原魔方：
Dactyl机器人



击败Dota2游戏国际职业选手：OpenAI Five



强化学习算法游戏模拟平台：Gym



通用强化学习算法测评平台：Universe

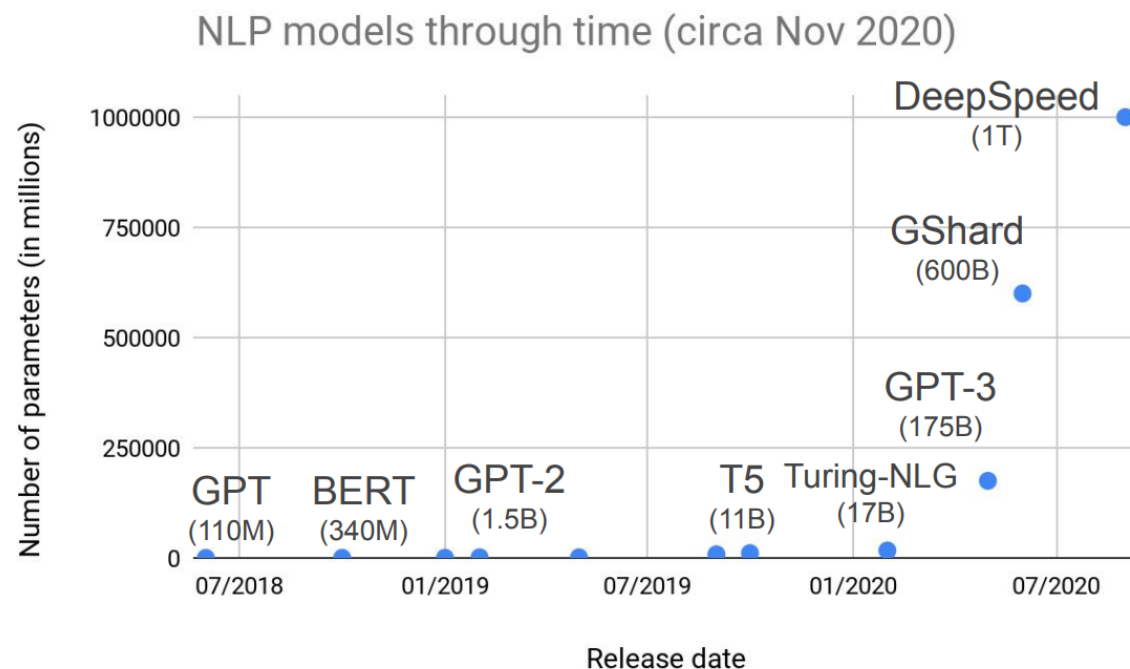
第一个通用预训练语言模型，启发了BERT，当前参数最大的模型，在多项NLP任务中表现优异

模型	总参数量	网络层数	训练数据量级
GPT	1.25 亿	12	十亿词
GPT-2	15 亿	48	百亿词
GPT-3	1750 亿	96	千亿词

大规模预训练模型的发展

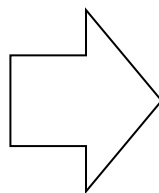
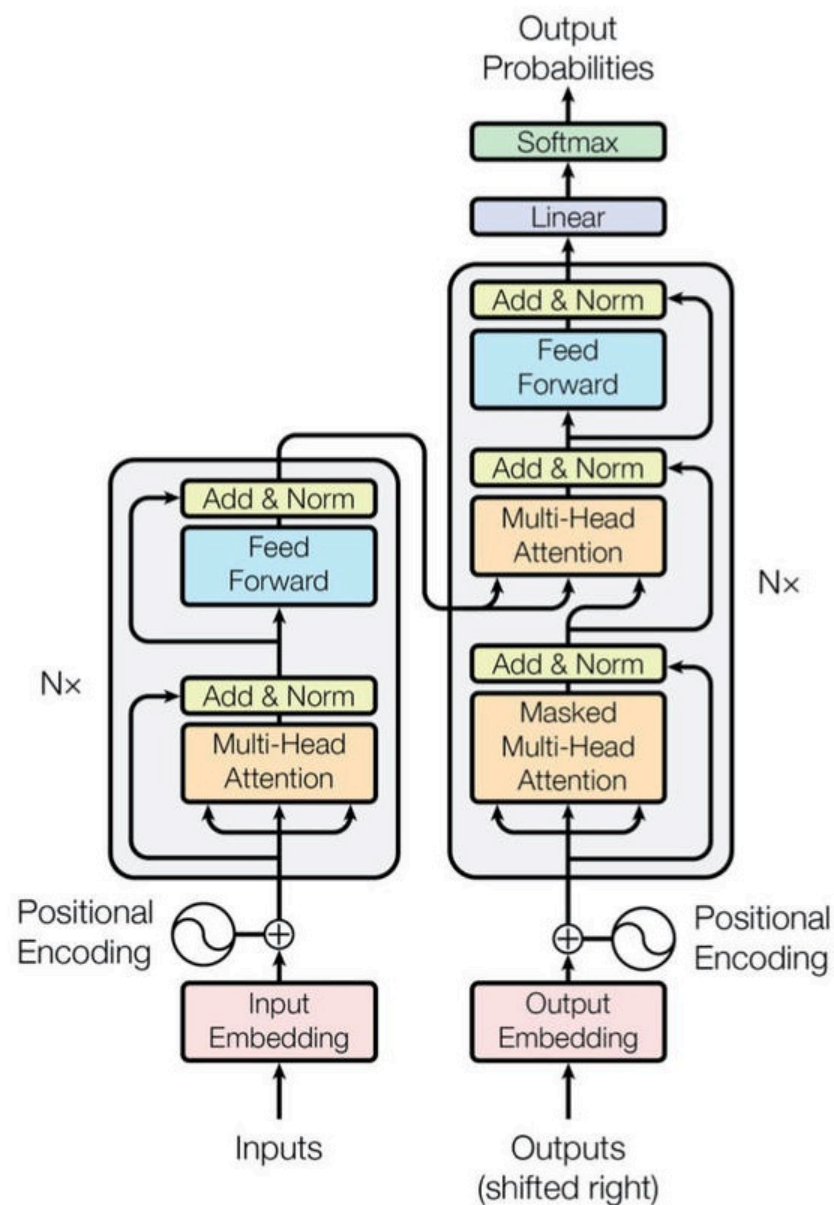
预训练模型规模以**每年约10倍**的速度增长，模型的通用智能水平显著增强

时间	机构	模型名称	模型规模	计算时间
2018.06	OpenAI	GPT	110M	3天
2018.10	Google	BERT	330M	50天
2019.02	OpenAI	GPT-2	1.5B	200天
2019.07	Facebook	RoBERTa	3.3B	3年
2019.10	Google	T5	11B	66年
2020.06	OpenAI	GPT-3	175B	355年



注：M-百万，B-十亿，最后一列计算时间为使用单块 NVIDIA V100 GPU训练的估计时间

GPT-3的阿喀琉斯之踵



Q: How many eyes does a giraffe have?

A: A giraffe has two eyes.

Q: How many eyes does my foot have?

A: Your foot has two eyes.

Q: How many eyes does a spider have?

A: A spider has eight eyes.

Q: How many eyes does the sun have?

A: The sun has one eye.

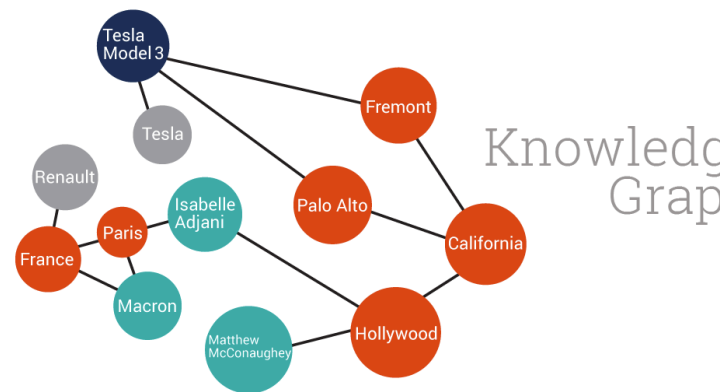
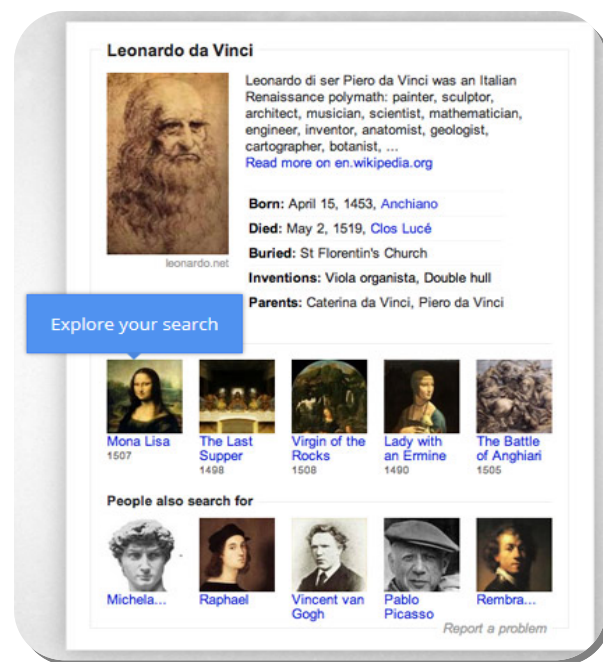
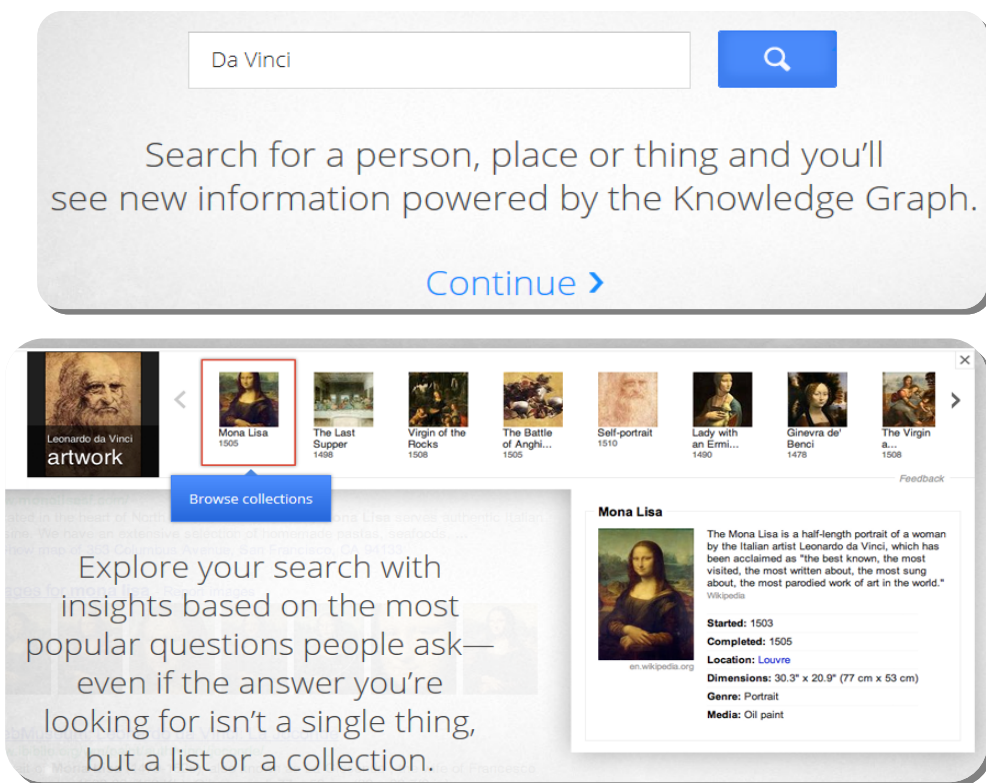
Q: How many eyes does a blade of grass have?

A: A blade of grass has one eye.

知识图谱

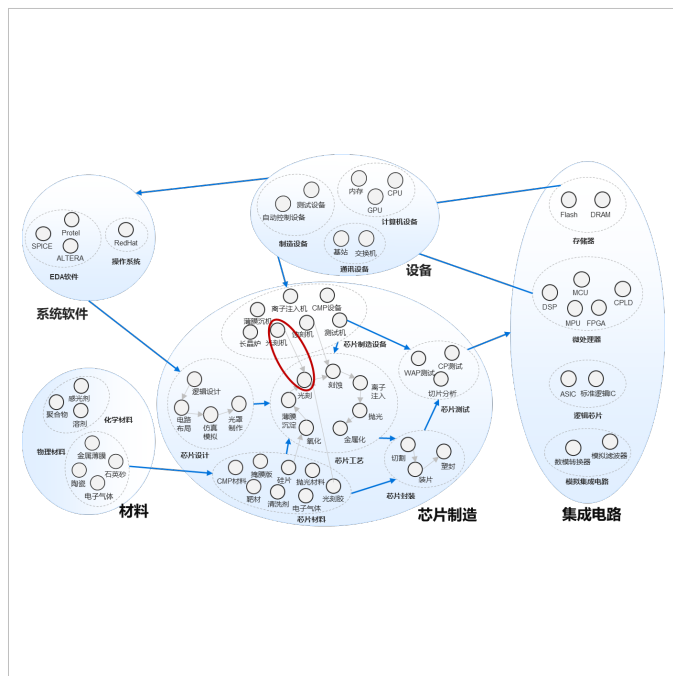
- 谷歌知识图谱

- “让搜索更智能”，与2012年5月16日加入谷歌搜索
- 最早的谷歌知识图谱源自 *CIA World Factbook*、*Freebase*、*Wikipedia* 的融合
- 包含：5亿对象、35亿实例和关系



知识图谱的先天缺陷

自我悖论：手工构建**成本高**和自动构建**精度低**



CYC是最早的知识图谱之一，每条知识断言手工成本为 **5.7美元***

— Douglas Lenat (CYC创始人)

NELL利用机器学习实现知识图谱自动构建，但**错误率增加10倍**

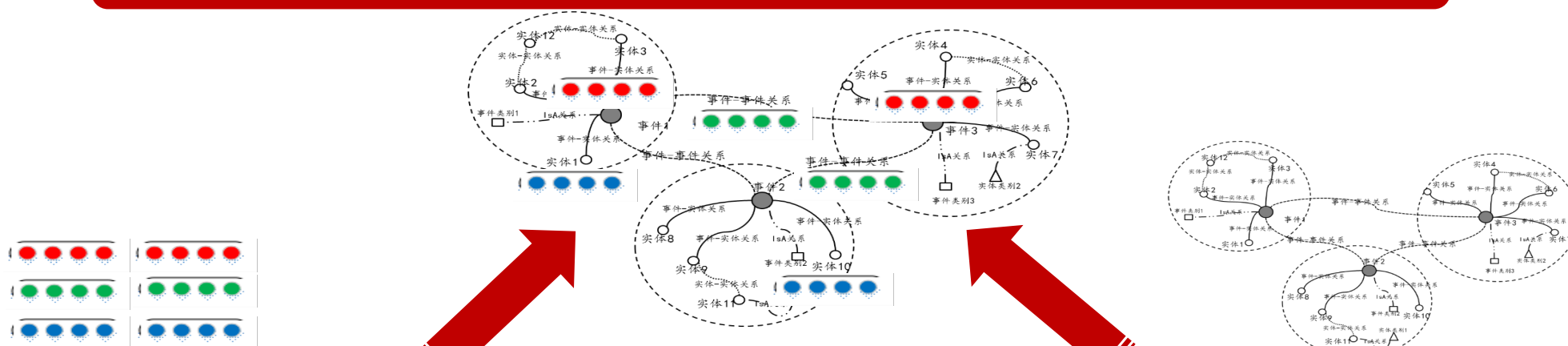
— Tom M. Mitchell (CMU教授，美工程院院士)

目前这两个项目都处于半停滞状态

* CYC有**2100万条**知识断言；而更大的知识图谱Freebase包含**30亿条**

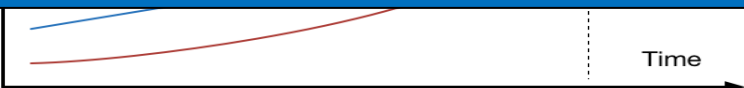
从计算模型看认知AI

基于感知的深度学习和基于认知的符号计算的融合



数据驱动

- 深度学习
- 举一反三
- 学习与记忆
- 关联挖掘
- 超大规模数据的预训练



Time

知识驱动

- 符号计算
- 举一反三
- 学习与理解
- 因果推理
- 常识知识驱动的可解释性



Time

知识 = 数值 + 符号

超越GPT-3的预训练

支持多种语言

同时支持中英文
基于至少数百GB的多语言
数据

在标准NLP任务上取得提升

认知能力突破（ e.g. 开放对话）

持续学习的能力

从经验中学习不断提升、记忆自身推理得出的结论

完成复杂任务的能力

行动和决策能力（e.g. 虚拟助手）、快速记忆知识和推理（e.g. 阅读文档并根据需求写代码）

模型架构

长距离序列建模能力
可以快速适应、可读写的记忆模块
对知识和推理进行模块化解耦

目标函数

同时具备双向建模能力和生成能力
从人类偏好中进行学习
将Meta Learning OOD泛化作为目标

让机器有“创造”能力

Context: 冬天, 中国哪座城市最适合避寒? 当然是海南三亚了。

CogGraph: 为什么? 因为三亚是中国的热带风景植物园。中国热带风景植物园位于三亚市北部的鹿回头山顶上, 占地1335亩, 距市区20公里, 与东线高速公路毗邻。园内有热带树木70科, 129属, 2300余种, 十万株。以热带雨林、草本花卉、棕榈植物为主, 还有仙人掌、兰花、爬藤等植物, 珠帘珍奇, 异彩纷呈。据园艺专家介绍, 植物园个季节均有专供盆景爱好者观赏的各式盆景, 花木品种多达500多种。主要摆设在东岸的滨湖和运动区, 水域面积有13700平方米, 有9000平方米的海上游乐设施, 还有长达1.5公里的健身步行道和水上活动区。在炎炎烈日之下, 到这里高大的棕榈林可即时披上如草帽般的轻纱, 与森林和风声融为一体, 增加观赏的乐趣。

V.S.

Context: 丁铭是清华大学计算机系的博士生, 导师是唐杰教授。

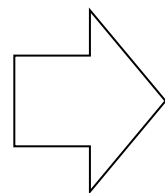
GPT2: 网名: 丁铭 性别: 男 出生年月: 1986年10月 籍贯: 河北省秦皇岛市 爱好: 广泛 学历: 博士 毕业院校: 清华大学 专业: 计算机应用 性别: 男 出生年月: 1986年10月 籍贯: 河北省秦皇岛市 爱好: 广泛 学历: 博士 毕业院校: 清华大学 专业: 计算机及其应用 性别: 男 出生年月: 1986年10月 籍贯: 河北省秦皇岛市 爱好: 广泛 学历: 博士 毕业院校: 清华大学 专业: 计算机应用 丁铭, 北京万维网拓科技有限公司董事长。北京万维网拓科技有限公司成立于2006年, 致力于互联网应用产品的研发、推广及服务。

让机器有“创造”能力

kamilan卡米兰2017秋季新款印花衬衫女宽松中长款长袖衬衣上衣



原图



自动生成

我们离AGI还差什么：更高级的认知能力

支持多种语言

同时支持中英文
基于至少数百GB的多语言
数据

在标准NLP任务上取得提升

认知能力突破（ e.g. 开放对话）

持续学习的能力

从经验中学习不断提升、记忆自身推理得出的结论

完成复杂任务的能力

行动和决策能力（e.g. 虚拟助手）、快速记忆知识和推理（e.g. 阅读文档并根据需求写代码）

模型架构

长距离序列建模能力
可以快速适应、可读写的记忆模块
对知识和推理进行模块化解耦

目标函数

同时具备双向建模能力和生成能力
从人类偏好中进行学习
将Meta Learning OOD泛化作为目标

认知AI的“九准则”

认知AI不仅会使用数据驱动的方法来建构更好的AI系统，还应该能将行为与意识联系起来，主动“学习”与创造行为

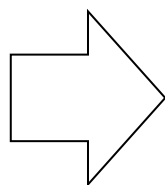
感知AI

数据驱动->模型训练

知识图谱->知识匹配

半监督学习->自监督学习

用户交互->强化学习



认知AI

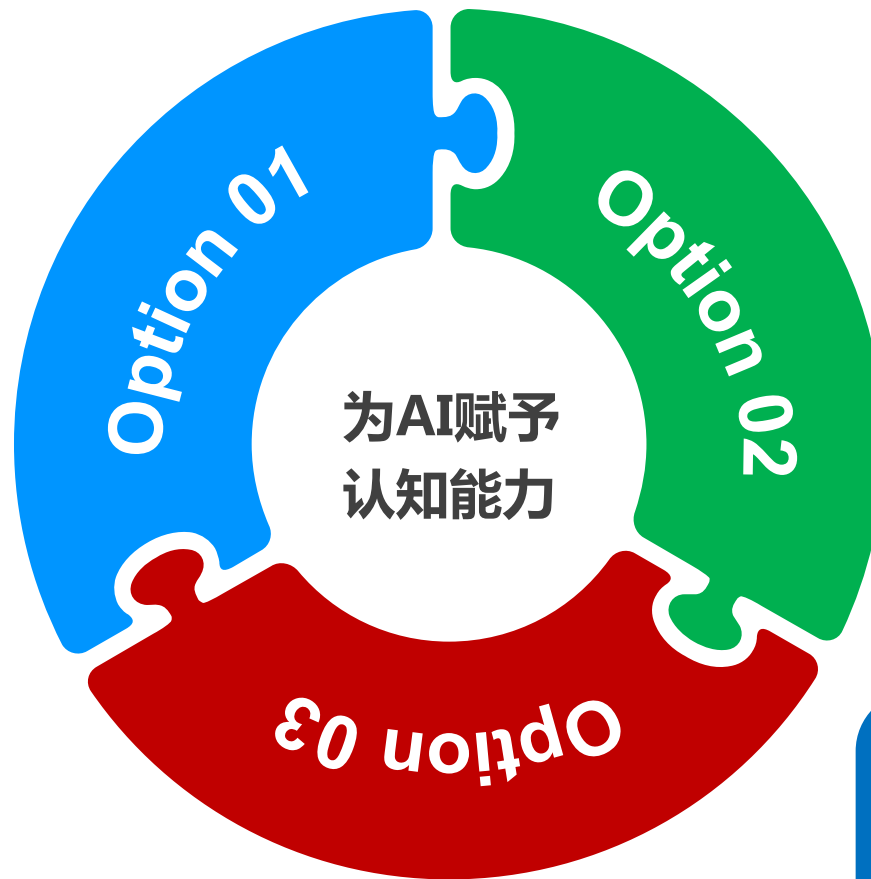
1. 适应与学习能力
2. 定义与语境化能力
3. 自我系统的准入能力
4. 优先级与访问控制能力
5. 召集与控制能力
6. 决策与执行能力
7. 错误探测与编辑能力
8. 反思与自我监控能力
9. 条理与灵活性之间的能力

认知图谱(Cognitive Graph)

常识图谱

Knowledge Graph

- 高精度知识图谱构建工具
- 领域知识图谱应用系统
- 超大规模常识知识图谱
- 基于知识图谱的推荐/搜索



认知推理

Cognitive Reasoning

- 超大规模预训练模型
- 内容自动生成
- “数字人”系统

逻辑生成

Logic Generation

- 突破系统2的认知推理
- 异常检测、关键线索挖掘
- 追根溯源、趋势预见

从脑认知看机器学习

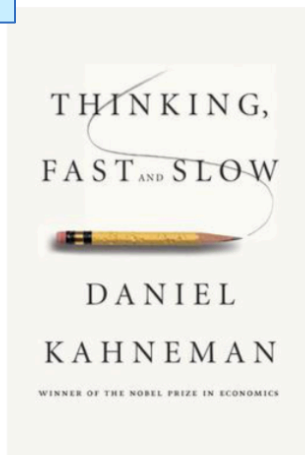
双过程理论 (认知科学) : 系统1vs. 系统2认知

当前深度学习仅解决了认知系统1的问题

System 1

- Intuitive, fast, **UNCONSCIOUS**, non-linguistic, habitual
- Current DL

直觉认知



System 2

- Slow, logical, sequential, **CONSCIOUS**, linguistic, algorithmic, planning, reasoning
- Future DL

逻辑认知



Manipulates high-level / semantic concepts, which can be recombined combinatorially

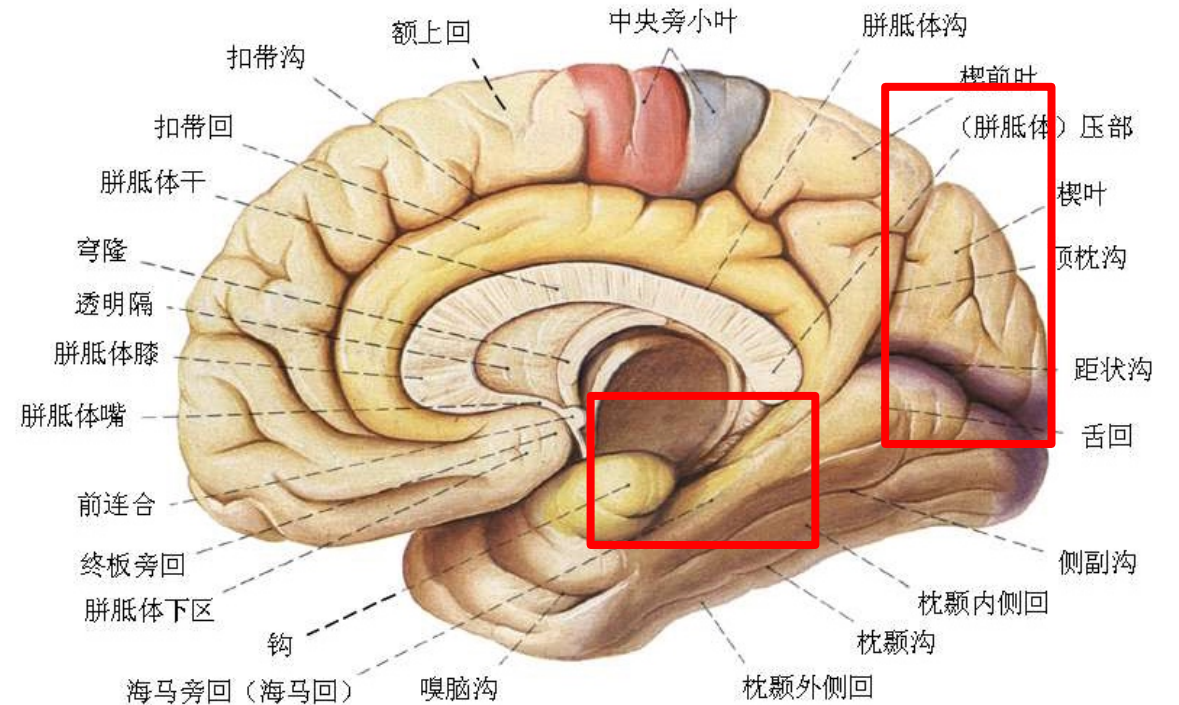
脑科学与认知智能

• 记忆

- 短期记忆在海马体转化为长期（陈述性）记忆
- 蛋白质磷酸化修饰（中期）
- 长时程增强、新蛋白产生（长期记忆）

• 推理

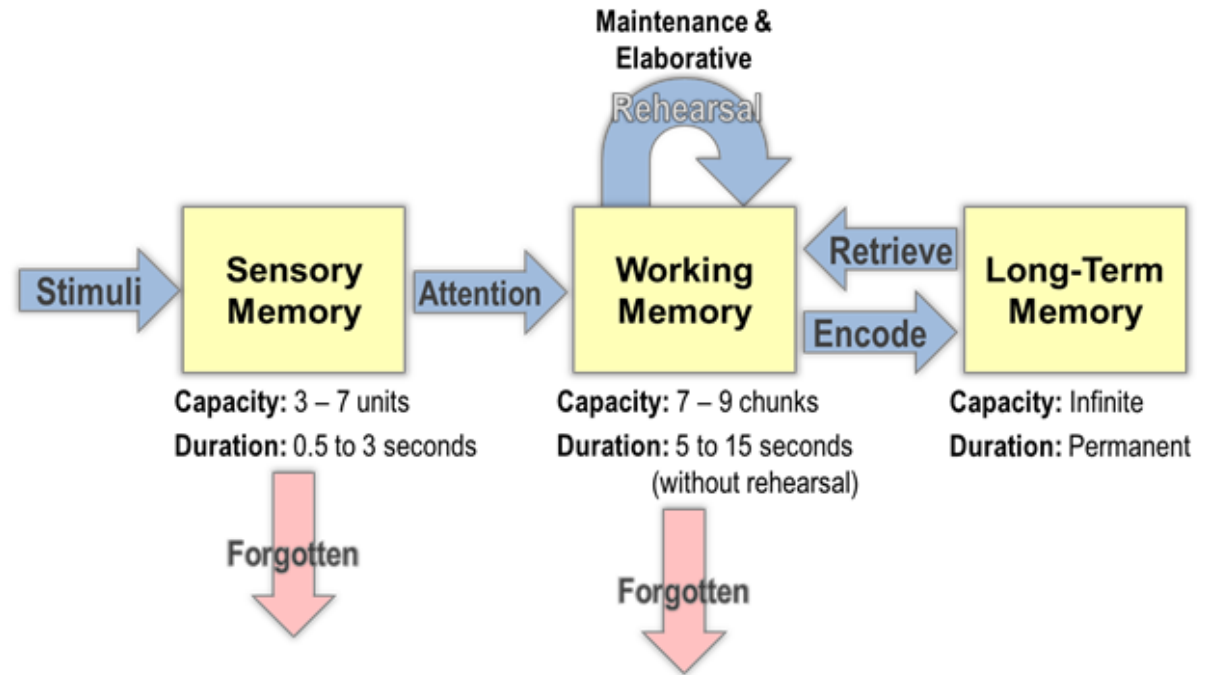
- 主要在前额叶中进行推理
- 语言脑区参与
- 对应工作记忆



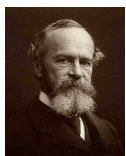
（睡眠时前额叶兴奋可以做能推断出自己在做梦的“清醒梦”）

记忆：工作记忆理论

- 尽管对于认知的微观机理尚未研究清楚，我们仍可以探究宏观框架。
- 巴德利的工作记忆（Working Memory）机制是里程碑式的工作，探究工作记忆调用多模态信息与长短期记忆转化（科万的分层注意力理论）。
- 全局工作空间理论（Global Workspace Theory）是巴斯等人对工作记忆模型的发展，认为“意识”是不同进程争夺全局空间传播信息的结果。



认知图谱¹



认知心理

提出认知的双通道理论，
美国心理学之父

William James等人



深度学习

以深度学习为代表的机器学习算法

Hinton等人

1842年

1979、1990年

1957、1988、2006年



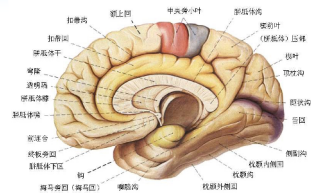
知识图谱

知识库之父、知识库问答，
语义网、知识推理

E. Feigenbaum、
T. Berners Lee

融合知识、学习与推理的
新一代认知引擎

1. 团队2018首次提出



认知图谱：面向认知的AI架构

超大规模预训练

错误探测与反馈
模型蒸馏与抽象

记忆模型

认知推理

“无意识”
探测

记忆模型1：

- artificial intelligence
- perceptron
- expert system

记忆模型2：

- data mining
- knowledge discovery
- association rule

记忆模型3：

- machine learning
- deep learning
- neural network

记忆模型4：

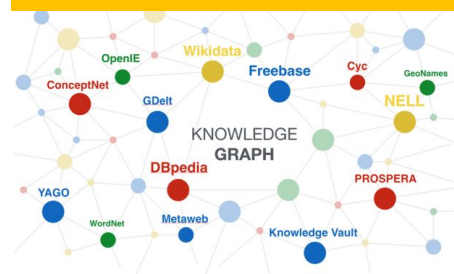
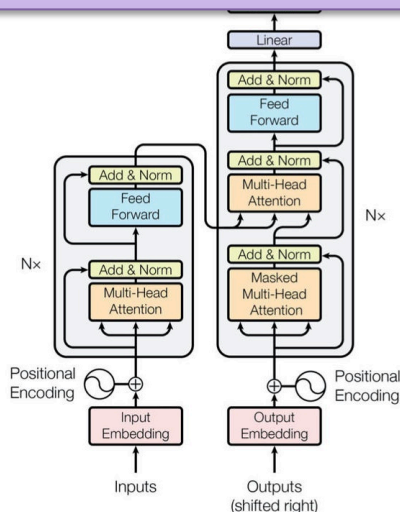
- machine learning
- deep learning
- neural network

自定义
条理与逻辑

“有意识” 决策

反馈与反思

查询接口



认知图谱智能服务平台

科技情报

安全

工业制造

金融

智慧教育

游戏

云服务

私有部署

联合研发

实时知识
抽取系统

面向知识计算与认知推理的智能服务

知识图谱

概念关系

知识获取

预训练

知识表示

生成模型

推理

认知推理

自校验

预测

新知识学习

趋势预测

决策

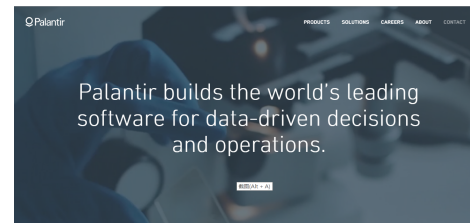
专家智库

智能问答

智能
总线

分布式知识存储、管理和计算平台

打造知识与认知推理
驱动的Palantir



Related Publications

For more, check <http://keg.cs.tsinghua.edu.cn/jietang>

- Wenzheng Feng, Jie Zhang, Yuxiao Dong, Yu Han, Huanbo Luan, Qian Xu, Qiang Yang, Evgeny Kharlamov, and Jie Tang. Graph Random Neural Networks for Semi-Supervised Learning on Graphs. **NeurIPS'20**.
- Ming Ding, Chang Zhou, Hongxia Yang, and Jie Tang. CogLTX: Applying BERT to Long Texts. **NeurIPS'20**.
- Jiezhong Qiu, Chi Wang, Ben Liao, Richard Peng, and Jie Tang. Concentration Bounds for Co-occurrence Matrices of Markov Chains. **NeurIPS'20**.
- Xiao Liu, Fanjin Zhang, Zhenyu Hou, Li Mian, Zhaoyu Wang, Jing Zhang, and Jie Tang. Self-supervised Learning: Generative or Contrastive. <https://arxiv.org/pdf/2006.08218.pdf>
- Jiezhong Qiu, Qibin Chen, Yuxiao Dong, Jing Zhang, Hongxia Yang, Ming Ding, Kuansan Wang, and Jie Tang. GCC: Graph Contrastive Coding for Structural Graph Representation Pre-Training. **KDD'20**.
- Zhen Yang, Ming Ding, Chang Zhou, Hongxia Yang, Jingren Zhou, and Jie Tang. Understanding Negative Sampling in Graph Representation Learning. **KDD'20**.
- Yukuo Cen, Jianwei Zhang, Xu Zou, Chang Zhou, Hongxia Yang, and Jie Tang. Controllable Multi-Interest Framework for Recommendation. **KDD'20**.
- Yuxiao Dong, Ziniu Hu, Kuansan Wang, Yizhou Sun and Jie Tang. Heterogeneous Network Representation Learning. **IJCAI'20**.
- Ming Ding, Chang Zhou, Qibin Chen, Hongxia Yang, and Jie Tang. Cognitive Graph for Multi-Hop Reading Comprehension at Scale. **ACL'19**.
- Jie Zhang, Yuxiao Dong, Yan Wang, Jie Tang, and Ming Ding. ProNE: Fast and Scalable Network Representation Learning. **IJCAI'19**.
- Yukuo Cen, Xu Zou, Jianwei Zhang, Hongxia Yang, Jingren Zhou and Jie Tang. Representation Learning for Attributed Multiplex Heterogeneous Network. **KDD'19**.
- Fanjin Zhang, Xiao Liu, Jie Tang, Yuxiao Dong, Peiran Yao, Jie Zhang, Xiaotao Gu, Yan Wang, Bin Shao, Rui Li, and Kuansan Wang. OAG: Toward Linking Large-scale Heterogeneous Entity Graphs. **KDD'19**.
- Qibin Chen, Junyang Lin, Yichang Zhang, Hongxia Yang, Jingren Zhou and Jie Tang. Towards Knowledge-Based Personalized Product Description Generation in E-commerce. **KDD'19**.
- Yifeng Zhao, Xiangwei Wang, Hongxia Yang, Le Song, and Jie Tang. Large Scale Evolving Graphs with Burst Detection. **IJCAI'19**.
- Yu Han, Jie Tang, and Qian Chen. Network Embedding under Partial Monitoring for Evolving Networks. **IJCAI'19**.
- Yifeng Zhao, Xiangwei Wang, Hongxia Yang, Le Song, and Jie Tang. Large Scale Evolving Graphs with Burst Detection. **IJCAI'19**.
- Jiezhong Qiu, Yuxiao Dong, Hao Ma, Jian Li, Chi Wang, Kuansan Wang, and Jie Tang. NetSMF: Large-Scale Network Embedding as Sparse Matrix Factorization. **WWW'19**.
- Jiezhong Qiu, Jian Tang, Hao Ma, Yuxiao Dong, Kuansan Wang, and Jie Tang. DeepInf: Modeling Influence Locality in Large Social Networks. **KDD'18**.
- Jiezhong Qiu, Yuxiao Dong, Hao Ma, Jian Li, Kuansan Wang, and Jie Tang. Network Embedding as Matrix Factorization: Unifying DeepWalk, LINE, PTE, and node2vec. **WSDM'18**.
- Jie Tang, Jing Zhang, Limin Yao, Juanzi Li, Li Zhang, and Zhong Su. ArnetMiner: Extraction and Mining of Academic Social Networks. **KDD'08**.



Thank you !

Jie Tang, KEG, Tsinghua U
Download all data & Codes

<http://keg.cs.tsinghua.edu.cn/jietang>
<https://keg.cs.tsinghua.edu.cn/cogdl/>
<https://github.com/THUDM>