

言語を獲得する知能の実現に向けた近年の取り組み

高木志郎



自己紹介

名前：高木志郎

専門：機械学習

連絡先：[@takagi_shiro](https://twitter.com/takagi_shiro)

研究のあり方に関心があり、個人で研究しています。



目次

1. DALL-E：言語AIの一つの到達点
2. DALL-Eに至るまで
3. 言語を獲得する知能の実現へ向けて（私的関心）
4. おまけ
5. 参考資料

* 技術的詳細には立ち入らずに大まかな流れだけをご紹介します

* Felix Hill “Why composition is DALL-E’s strength, not its weakness” に強い影響を受けてます

* 独断と偏見に塗れた内容ですが言語のAIの世界は広大なのでぜひ色々調べてみてください

1. DALL-E：言語AIの一つの到達点

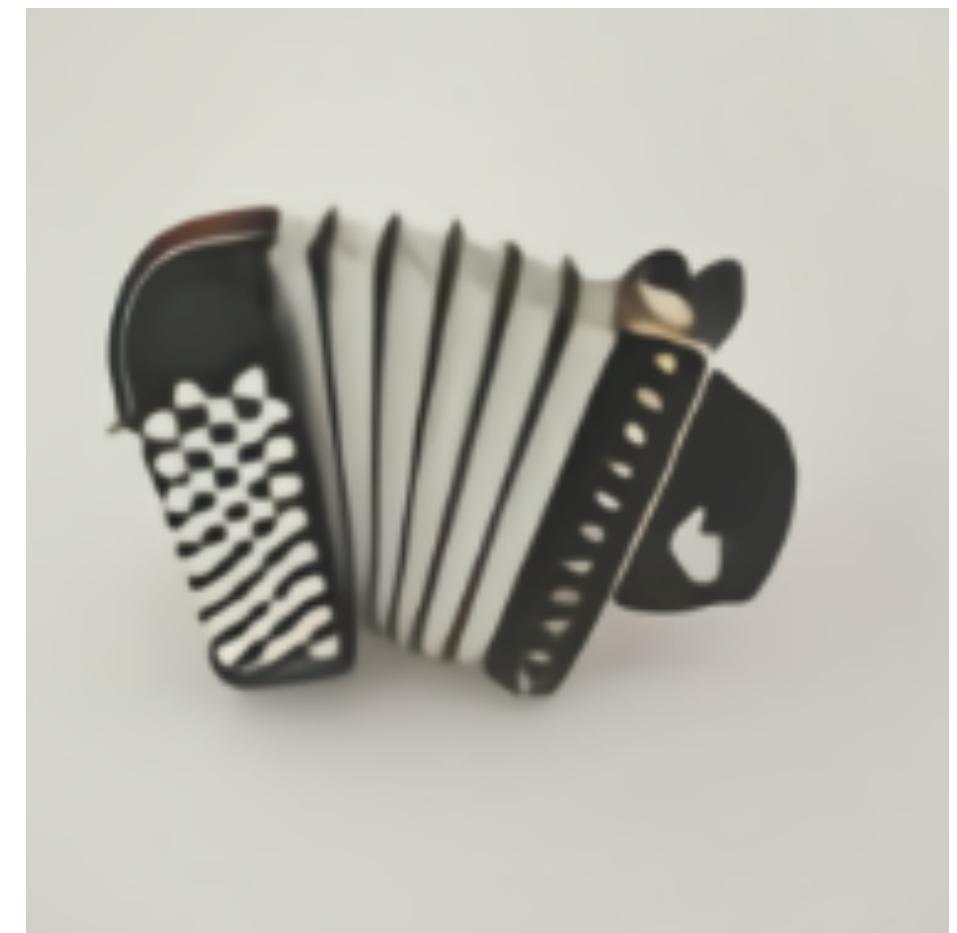
高木志郎



"An illustration of a baby daikon radish in a tutu walking a dog"



"An armchair in the shape of an avocado"



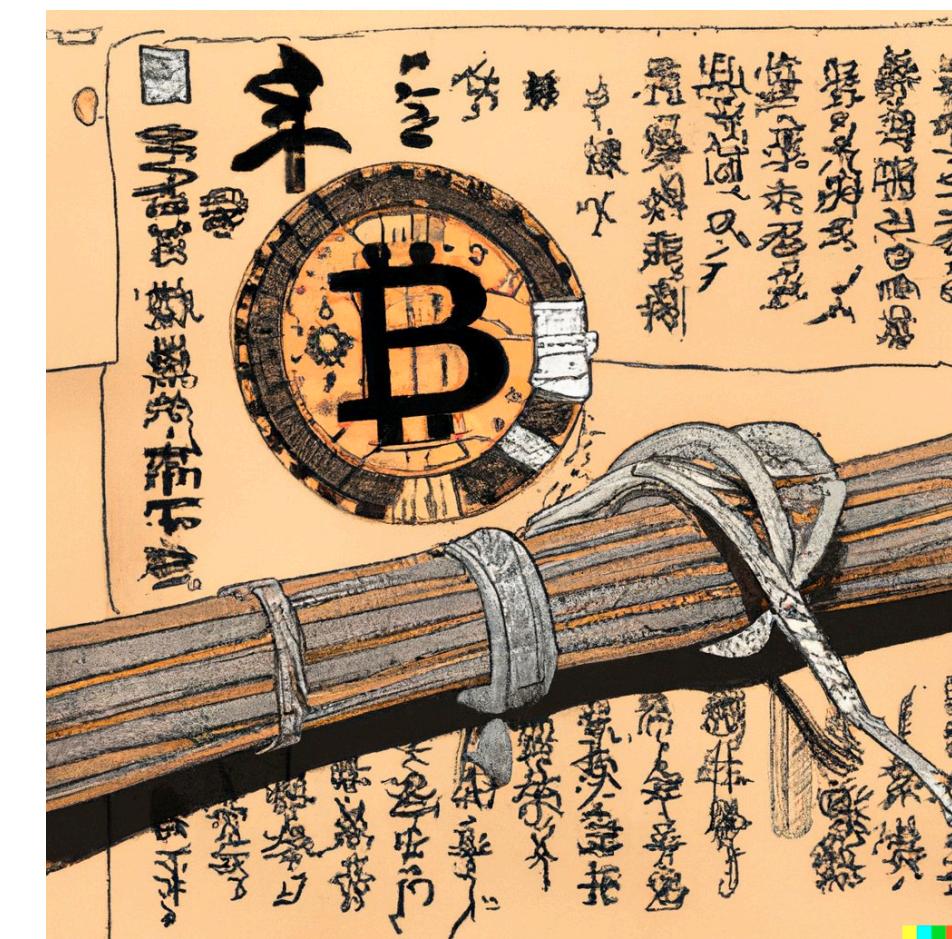
"A tapir made of accordions"



"A bowl of soup that looks like a monster knitted out of wool"



"An astronaut riding a horse in a photorealistic style"



"Pre-modern Japanese scroll about Bitcoin"

<https://openai.com/dall-e-2/>

<https://twitter.com/hardmaru/status/1517470815222386694>

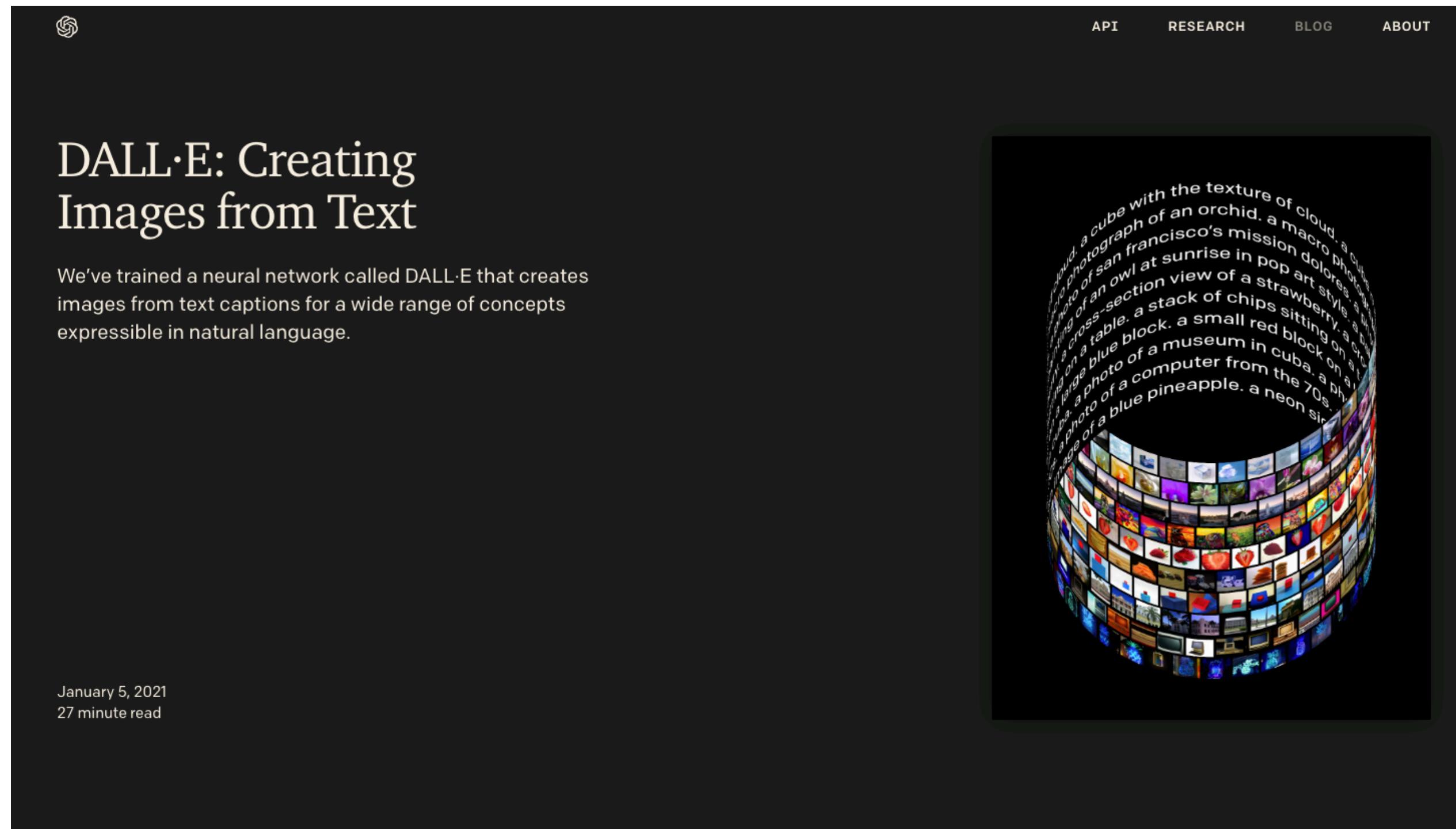


"Oriental water painting of Hello Kitty programming on a MacBook during the Song dynasty"

<https://twitter.com/hardmaru/status/1528019308307107840>

DALL·E (2)

汎用人工知能の実現を目指す会社であるOpenAIが2021年と2022年に生み出したAI
与えられた文章に対応する画像を生成することができる



DALL·Eの何がすごいのか

DALL·Eが**構成性 (compositionality*)** を獲得しているように見えるのが、すごい

構成性とは、**複数の要素を組み合わせて意味を構成する**という言語/思考の性質



"An armchair in the shape of an avocado"



* むしろDALL·Eはある種のcompositionalityは獲得できていないことが課題なのではないかとの声もあります

<https://www.avocadosfrommexico.jp/>
ditreitalia.com/en/products/armchairs/round

Compositionalityの何がすごいのか

構成性により**有限の要素から無数の意味を構成できる**のが言語のすごさの一つ (**productivity**)



"An armchair in the shape of an avocado"



"An armchair in the shape of a peach"

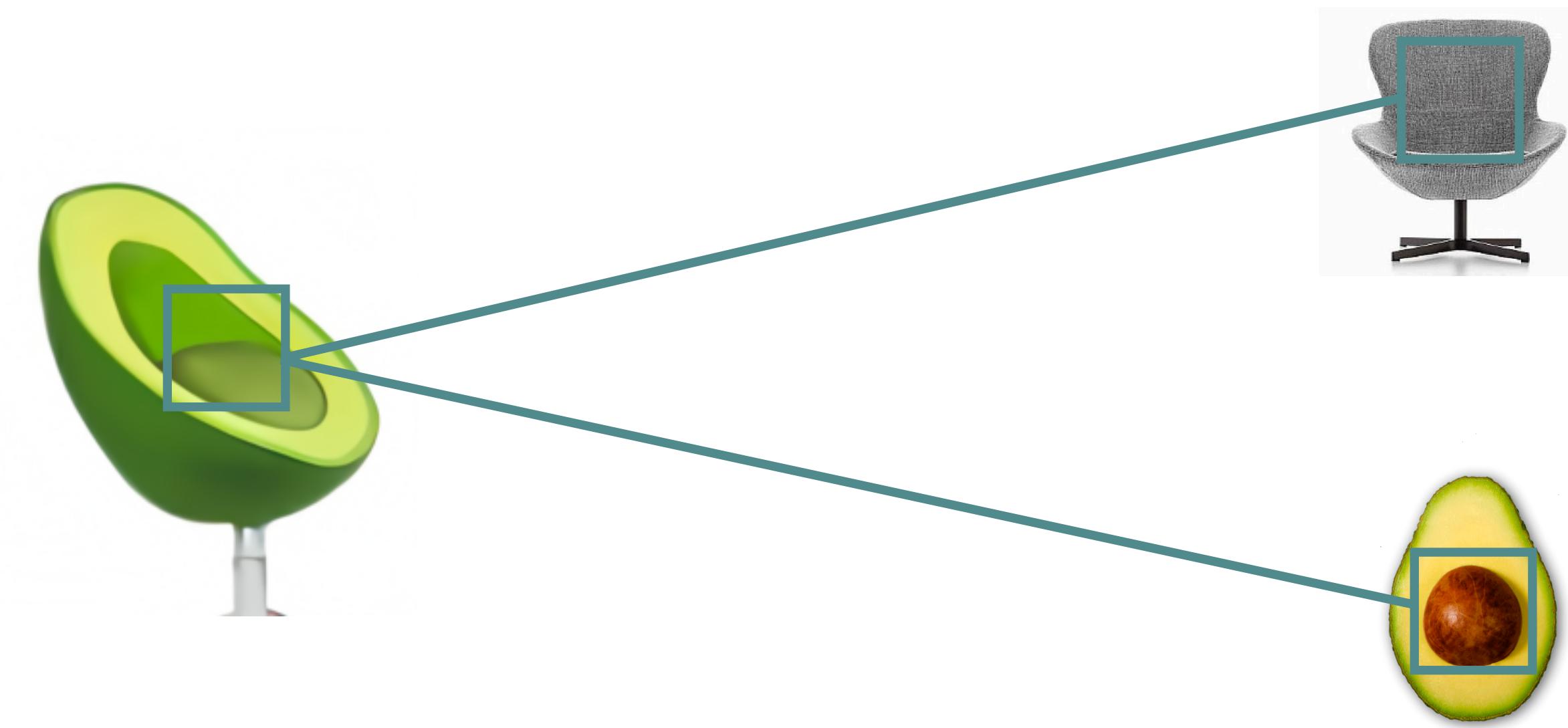


"An armchair in the shape of a brain coral"

...

Compositionalityが獲得できたことの何がすごいのか

構成の仕方は無数であるにも関わらず、「人間にとて自然な」ものを生成できてるのがすごい
それを人に教わるでも世界から学ぶでもなく画像と言語のデータのみから獲得するのがすごい
画像と言語のみから人間社会における構成の仕方の暗黙的なルールを獲得しているのがすごい



2. DALL-Eに至るまで

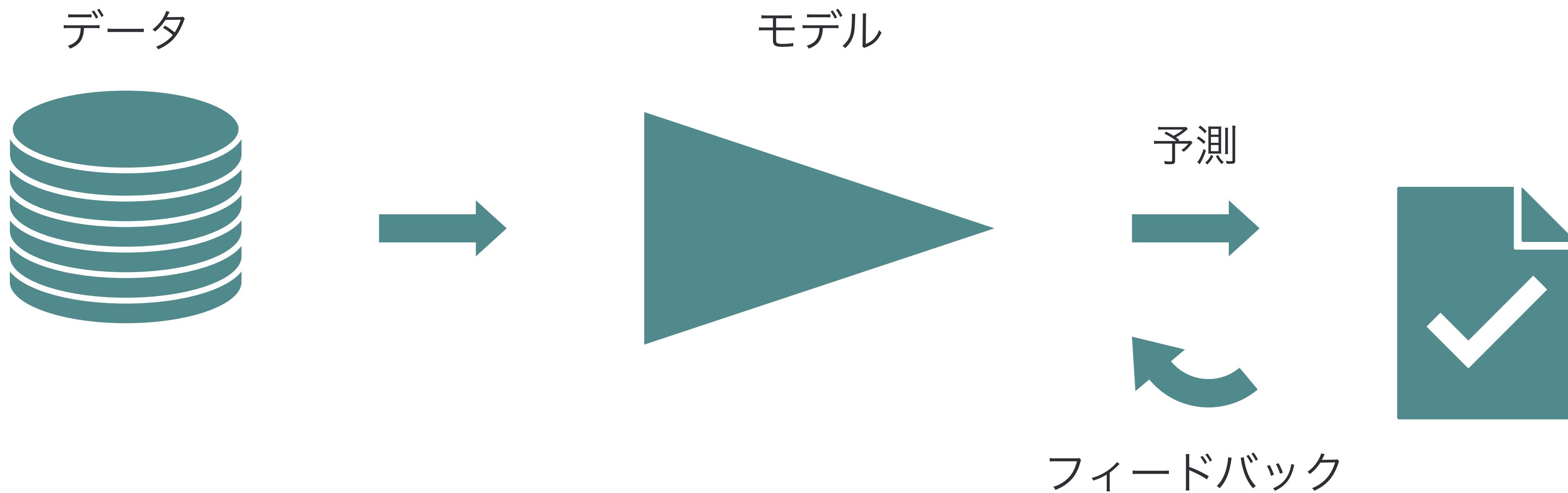
高木志郎

(主流の) 機械学習 (AI) の概要

パラメータ付き関数である「モデル」が「データ」から何らかの「予測」を出力する

ある基準に照らしてその予測に対する「フィードバック」が与えられる

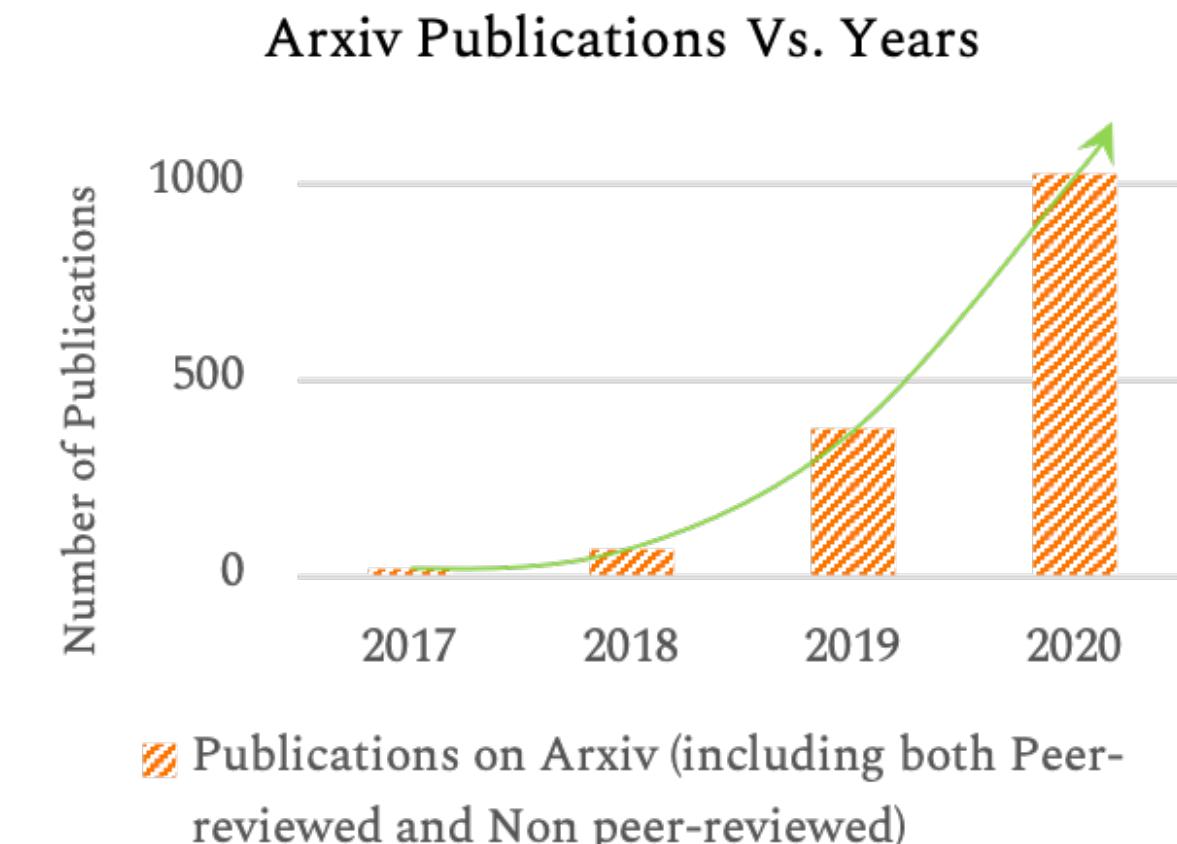
所定の「アルゴリズム」でパラメータを修正して、モデルを所望の関数に近づけることを目指す



Transformerの登場（2017年）

2017年にTransformerというモデルが提案された

Transformer登場以降自然言語処理のモデルはTransformer一色に



Attention is all you need

[A Vaswani, N Shazeer, N Parmar... - Advances in neural ... , 2017 - proceedings.neurips.cc](#)
... the number of **attention** heads and the **attention** key and value dimensions, keeping the amount of computation constant, as described in Section 3.2.2. While single-head **attention** is 0.9 ...
☆ 保存 囗 引用 引用元 42753 関連記事 全 35 バージョン ≫

Vaswani+ 2021 Attention is All You Need

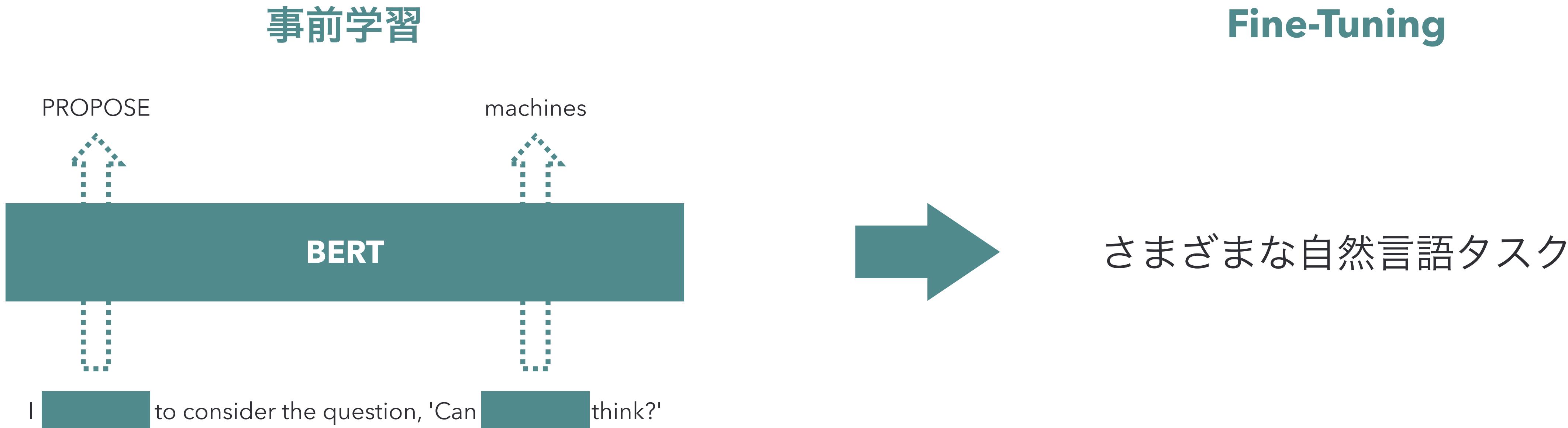
Khan+ 2021 Transformers in Vision: A Survey

https://scholar.google.co.jp/scholar?hl=ja&as_sdt=0%2C5&q=attention+is+all+you+need&btnG=&oq=attention

BERTの登場（2018年）

Transformerベースのモデルおよびその学習方法が提案された（BERT）

以降、大規模モデルを大規模データで事前学習し汎用言語モデル実現を目指す流れが加速



GPT-3の登場

1750億個のパラメータからなるTransformerベースのモデルであるGPT-3の登場

モデルに対して言語で指示を出す

言語に関わるあらゆるタスクを解いてしまうとして界隈に衝撃を与えた

コピーライトの生成

レシピの生成

コードの生成

COPYWRITING SUMMARIZATION PARSING UNSTRUCTURED TEXT CLASSIFICATION

Create promo copy for the FamilyTime mobile application. It allows unlimited uploading, special filters and makes it easy to create albums of photos and videos. It runs on iOS and Android:

FamilyTime is the perfect app to share and create memories with your family. FamilyTime is a free, unlimited photo and video editor that allows you to create albums of your memories. It has a special filter for the perfect summer look and is so simple to use.

Write a recipe based on these ingredients and instructions:

Frito Pie

Ingredients:

- Fritos
- Chilli
- Shredded cheddar cheese
- Sweet white or red onions
- Sour cream

Instructions:

- Preheat oven to 375 degrees F.
- Spread Fritos in an even layer on a baking sheet.
- Top with chilli and shredded cheese.
- Top with small onions.
- Bake in preheated oven for 15 minutes.
- Remove from oven and top with sour cream.
- Enjoy!

T TO SQL TRANSLATION CALLING AN API VIA NATURAL LANGUAGE CODE CONTINUATION

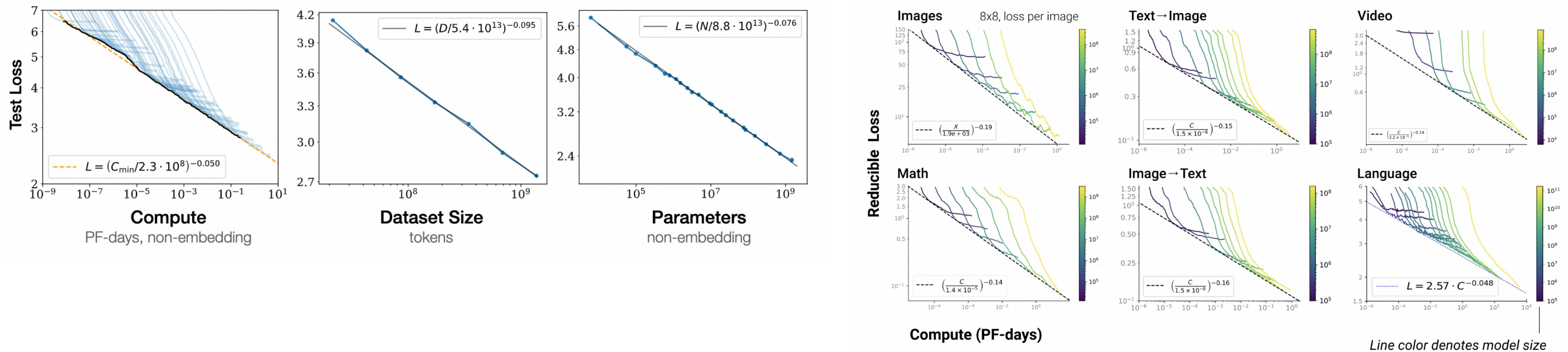
```
....  
Python 3  
Get the current value of a Bitcoin in US dollars using the  
bitcoinccharts api  
....  
  
import requests  
import json  
  
def get_bitcoin_price():  
    url = 'http://api.bitcoinccharts.com/v1/weighted_prices.json'  
    response = requests.get(url)  
    data = json.loads(response.text)  
    return data['USD']['7d']  
  
if __name__ == '__main__':  
    print(get_bitcoin_price())
```

More is Different

性能がデータ数・計算量・パラメータ数に対し幂で増加し現状上限が見られない (**Scaling Law**)

さらに、これは画像・言語・複数モダリティ・演算・動画などあらゆるドメインで成り立つ

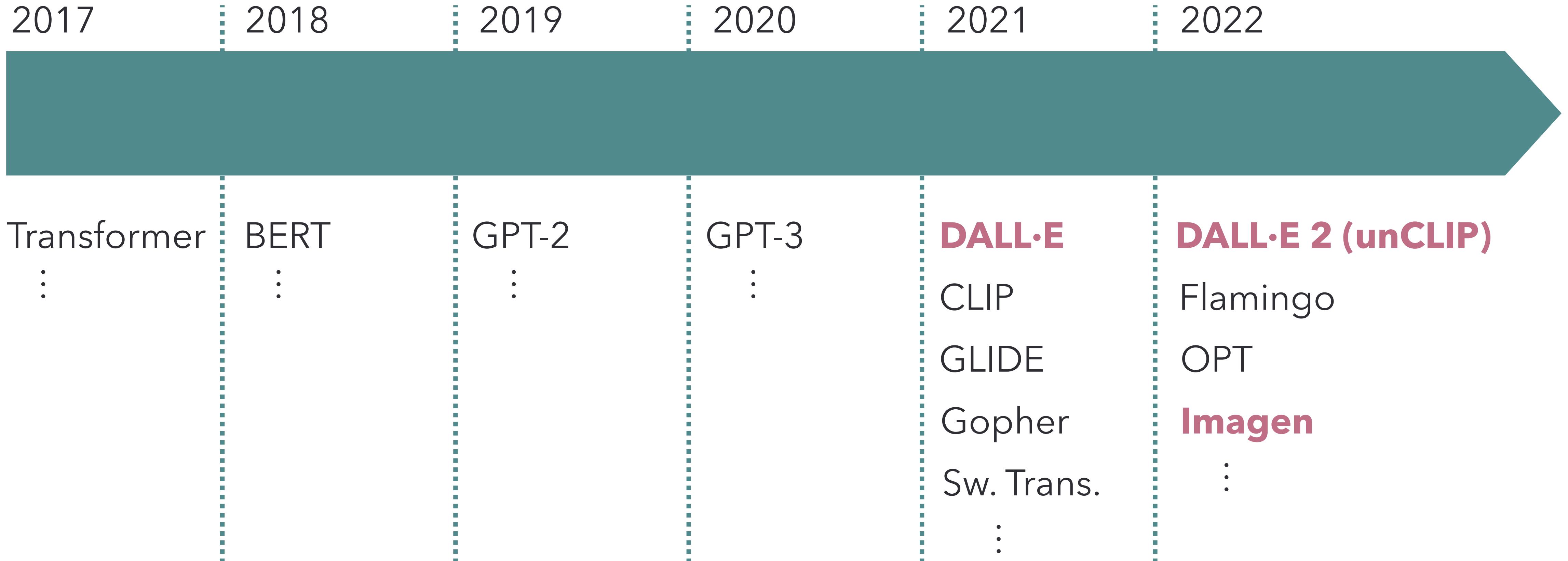
GPT-3の衝撃も相まって、**大企業主導で大規模化により人間のような人工知能を目指す動きが加速**



Kaplan+ 2021 Scaling Laws for Neural Language Models

Henighan+ 2021 Scaling Laws for Autoregressive Generative Modeling

DALL·Eまでの流れのまとめ



Imagen

Google発のモデルで、DALL·E 2登場(4/13)から1ヶ月(5/23)で早くも**DALL·E 2を超えた**との評価



A giant, stone hand gently balancing a car on the tip of the Eiffel Tower.



A cute corgi lives in a house made out of sushi.



A brain riding a rocketship heading towards the moon.



Teddy bears swimming at the Olympics 400m Butterfly event.

<https://imagen.research.google/>

https://twitter.com/Chitwan_Saharia/status/1529904034957643778

Saharia+ 2022 Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding

3. 言語を獲得する知能の実現へ向けて（私的関心）

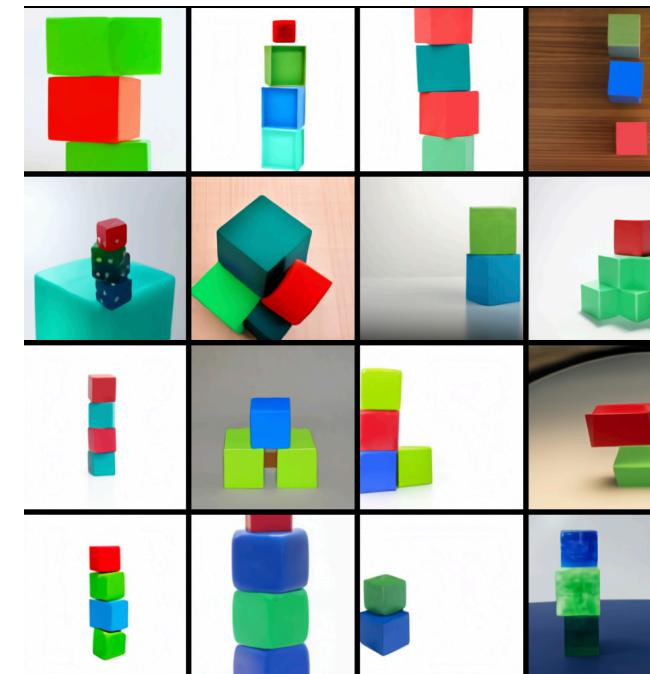
高木志郎

Systematicityが不完全

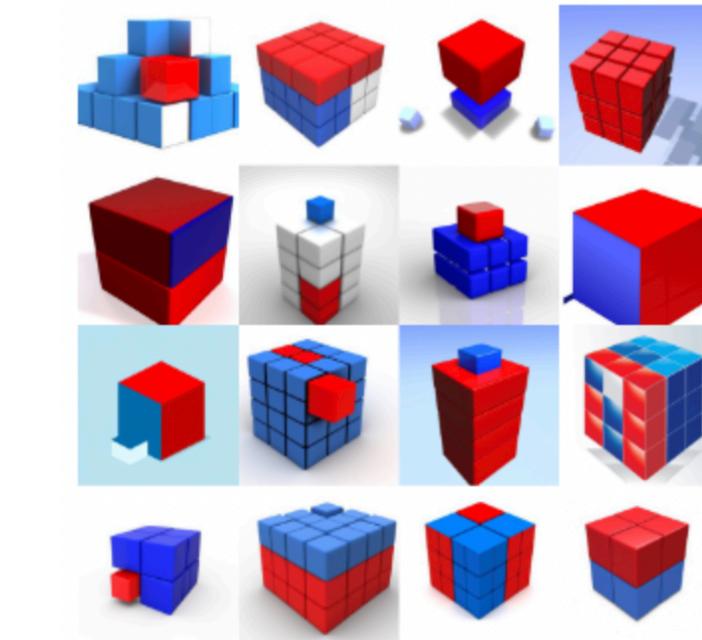
人はある文（思考）Aを理解できれば意味的に類似した文（思考）Bを理解可能（**Systematicity***）

DALL·Eはある種のsystematicityが苦手？（得意なcompositionalityとそうでないものがある）

言語を通じて暗黙的に獲得している世界・人間社会のルールをより厳密に獲得することが必要



"a stack of 3 cubes. a red cube is on the top, sitting on a green cube. the green cube is in the middle, sitting on a blue cube. the blue cube is on the bottom."



"Red cube on top of a blue cub"

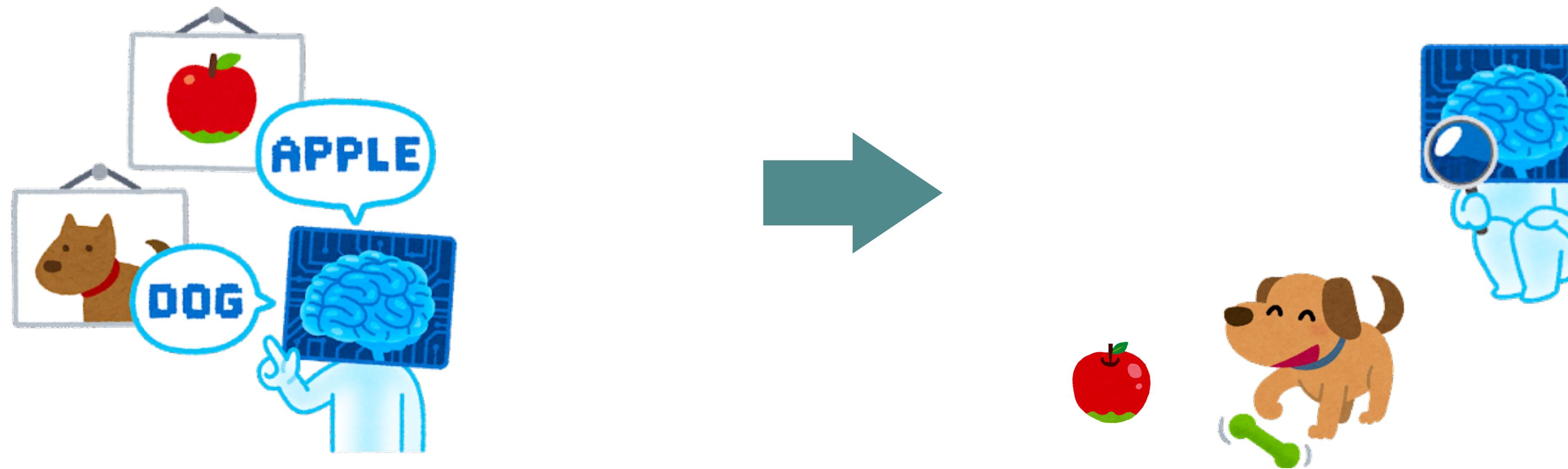
Embodied/grounded language learning

人間言語はこの世界で用いられるものとして存在 ("語の意味はその使用")

言語の構成則はこの世界のありように強く影響されている

より多様で豊かな身体と感覚器官を持ちより豊かな世界と対応づけて言語を学習することが重要

世界と対応づいた言語学習は今後の重要な方向性の一つ (**Grounded language learning**)



Human in the Loop

言語の構成則はあくまで人間にとて自然なものなるものが社会的に合意されているだけ

言語発達ではその言語を持つ社会の構成員との相互作用が必要 (**zone of proximal development**)

真に言語を獲得する知能の実現には人間を陽に組み込んだ学習が重要そう (**Human in the Loop**)

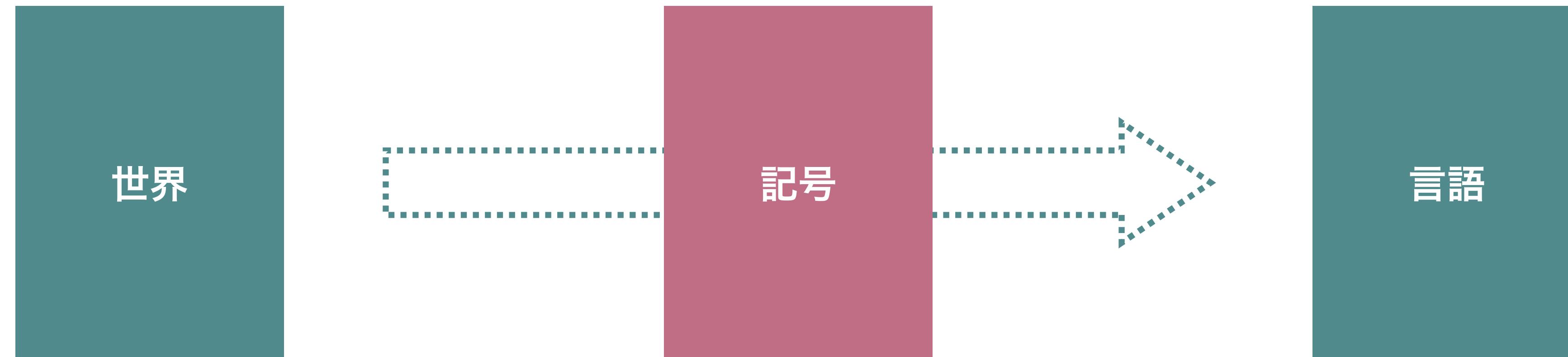
Symbol Emergence

これまでの議論はいずれも言語データありきの議論

人間が生物として言語を獲得した背景には言語獲得につながる前駆体の存在があると考えられる

例えば複雑な行動が単純な行動から構成される点で行動と言語は類似している (**action grammar**)

世界を素朴な記号の体系で表象する能力の探究は言語獲得を目指す上で重要 (**symbol emergence**)



Lashley 1951 The Problem of Serial Order in Behavior

Santoro+ 2021 Symbolic Behaviour in Artificial Intelligence

Taniguchi+ 2015 Symbol Emergence in Robotics: A Survey

まとめ

近年言語に関するAIは急速な発展を遂げている

これまで困難であった様々なことが克服されており、今後も一層の飛躍が見込まれる

言語を獲得する知能の実現には現実世界と人間社会に根差した言語学習が重要だと考えられる

分散表現によるボトムアップなデータからの帰納的な学習から、如何にして言語や記号の操作に至り、そして如何にして論理や数学という形式的で演繹的な手法に辿り着くのか、は昔からの課題であり人間のような知能を実現するに当たっての最も大きな課題の一つ

"It is obvious to many others besides ourselves that parallel processing and sequential symbol processing must somehow occur in the same system"

[McClelland & Rumelhart 1987]

論理/数学をご専門とされる方々も是非これらの課題へのご挑戦を！！

4. おまけ

Automated Theorem Proving (ATP) × Machine Learning (ML)

定理の自動証明 (ATP) に機械学習を用いる研究が近年注目を浴びる

$$\overbrace{\text{REWRITE_TAC}}^{\text{tactic name}} \quad \overbrace{[\text{PREMISE1} ; \text{PREMISE2}]}^{\text{list of premises}}$$

Premise selectionへの活用 [Urban+ 2008]

DeepMath: 深層学習を初めて premise selectionへ活用 [Irving+ 2016]

深層学習ベースの手法の台頭 (Gamepad, HOList, DeepHOL, CoqGym, GPT-f, TacticZero..)

[Rabe & Szegedy 2021] は人工數学者実現へ向けては自然言語による学習が重要だと主張

Towards the Automatic Mathematician

MLでATPおよび人工数学者を目指す試みは帰納と演繹を結ぶという意味でアツい
「形式」的操作とデータに潜む人間のある種の「直観」が交わるという意味でアツい
数学という営みに対する理解が深まる可能性があるという意味でアツい
研究という営みの自動化という意味でアツい
論理/数学をご専門とされる方々でご関心持たれた方是非当該分野へもご参入を！！

Conference on Artificial Intelligence and Theorem Proving (AITP)

人工知能/機械学習×定理自動証明な研究トピックを扱う国際会議

小規模だが分野の主要人物や重要な成果が集まるようなので、参加すると楽しいかも



7th Conference on Artificial Intelligence and Theorem Proving
AITP 2022
September 4-9, 2022, Aussois and online, France

Background
Large-scale semantic processing and strong computer assistance of mathematics and science is our inevitable future. New combinations of AI and reasoning methods and tools deployed over large mathematical and scientific corpora will be instrumental to this task. The AITP conference is the forum for discussing how to get there as soon as possible, and the force driving the progress towards that.

This year AITP will be co-located with a meeting of WG5 of Cost Action European Research Network on Formal Proofs.

Topics

- AI and big-data methods in theorem proving and mathematics
- Collaboration between automated and interactive theorem proving, in particular their AI/ML aspects
- Common-sense reasoning and reasoning in science
- Alignment and joint processing of formal, semi-formal, and informal libraries, Formal Abstracts
- Methods for large-scale computer understanding of mathematics and science
- Combinations of linguistic/learning-based and semantic/reasoning methods
- Formal verification of AI and machine learning algorithms, explainable AI

5. 參考資料

参考資料

Felix Hill "Why composition is DALL-E's strength, not its weakness"

Aurelie Herbelot "How to Stop Worrying About Compositionality"

Stanford Encyclopedia of Philosophy "Compositionality"

GPT-3の衝撃