Artificial Intelligence: Past, Present and Future

Shusaku Tsumoto

Dept of Medical Informatics, Faculty of Medicine, Shimane University

tsumoto@med.shimane-u.ac.jp

President of Japan AI Society

Outline

- What is Artificial Intelligence ?
- Four Classical Problems in AI
  - Role of Machine Learning
- Emergence of Generative AI
  - (Diffusion Model (Medical Image))
  - Chat GPT (Language)
- Future Research
  - AI for Data Preprocessing
- Summary

What is Artificial Intelligence?

- Goal: Implementation of intelligence into Computers
- Starts from the introduction of computers (in 1950’s): switch on-off ⇔ Boolean
- The word “Artificial Intelligence” is declared during Dartmouth Conf in 1956.

AI: booming?

- AI players beat the major champions in Chess, Shogi and Igo.
  - The point is:
    - High speed computation
    - Electronic playing records
- Generative AI (ChatGTP)
  - Various data stored as digital texts
  - Learning from digital texts
      - Large linguistic model (LLM)
- AI application areas are growing.

Past: Four Classical Problems

Four Classical Problems in AI

- Intelligent Game: Chess (Shogi, Igo)
- Problem Solving
  - Automated Diagnosis
- Machine Translation
  - Translation from Japanese to English
- Turing test (Interactive AI)

Intelligent Games

- Chess: beats the world champion in 1997.
  - Read ahead in chess moves deeper
    - Refining evaluation functions

    (Contribution of Machine Learning: 0%)
- Shogi: beats the champion in 2015.
  - Learning the evaluation function from records / Parallel Learning

    (Contribution of ML: 30%)
- Igo: 2017
  - Deep Learning

  (Contribution of ML: 100%)

Automated Diagnosis

- Implementation of reasoning of domain experts
- MYCIN: Diagnosis of bacterial infection (1973)
  - Rule-based system.
- The system can only diagnose typical cases and does not learn from failure.
  - → Machine Learning needed.
- Difficult to diagnose images, waveform and time series (multi-media).

MYCIN

- Shortliffe, Buchnan: Stanford University AI Group

- Rule-based diagnosis of bacterial infection
- Consists of 500 rules
- Accuracy: 65% (Domain experts: 80%)

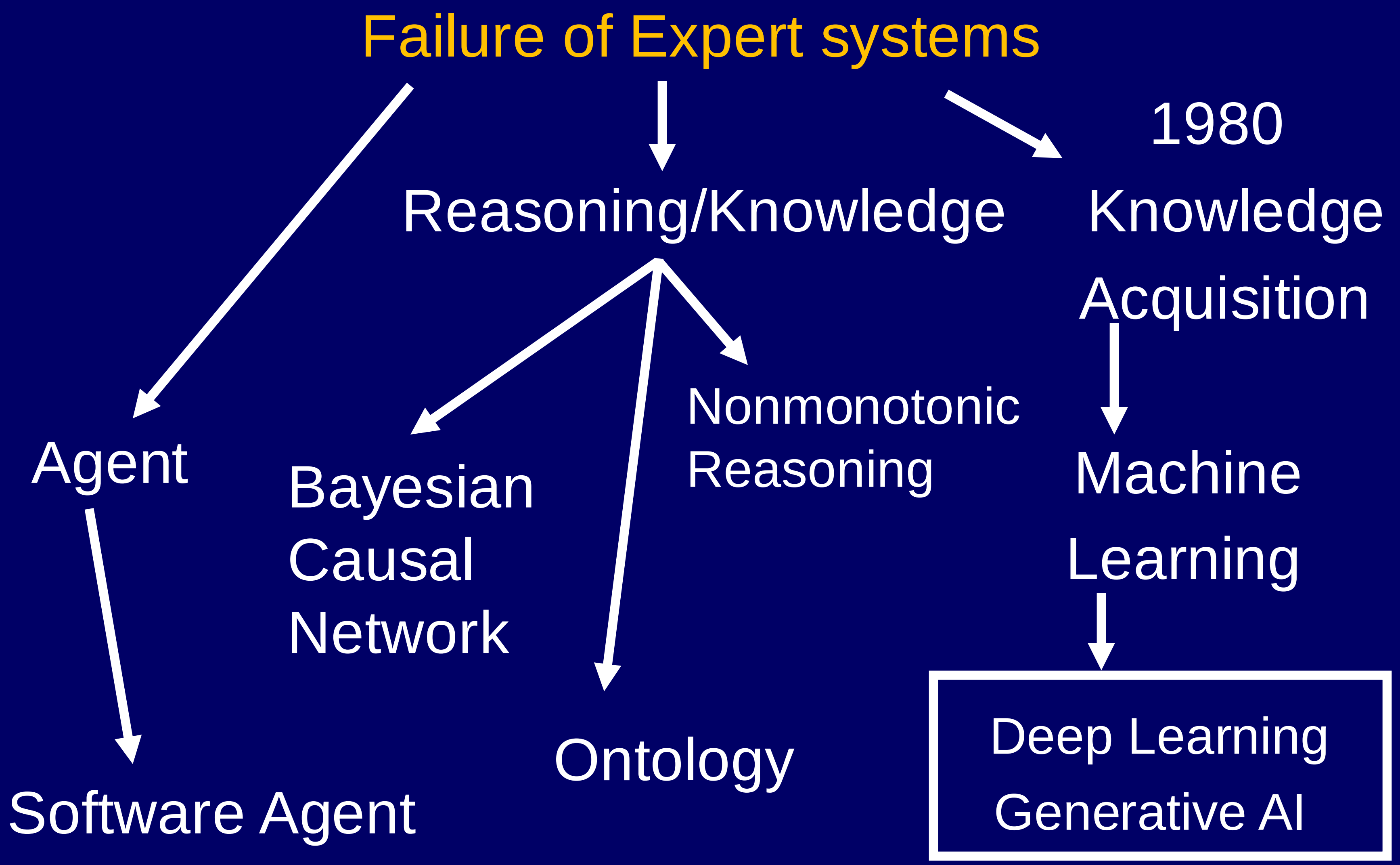
```
(defrule 165
  if (gram organism is pos)
     (morphology organism is coccus)
     (growth-conformation organism is chains)
  then .7
     (identity organism is streptococcus))
```

Limitations of Expert System ('80s)

- Redundant Inputs
- Can diagnose only typical cases
  - Difficult to diagnose complicated cases
- No support for whole decision processes
  - Diagnosis => Therapy
- Cannot learn from failures.
- Users learn the reasoning rules of the system.
  - Application domains: limited.

10

Progress in AI research



s

Present:

Machine Learning

Generative AI

AI Mapβ 2.0

Overview figure of issues and technologies for novice AI researchers as well as researchers and practitioners in other fields

AI Map Task force,
the Japanese Society for Artificial Intelligence

Table of Contents

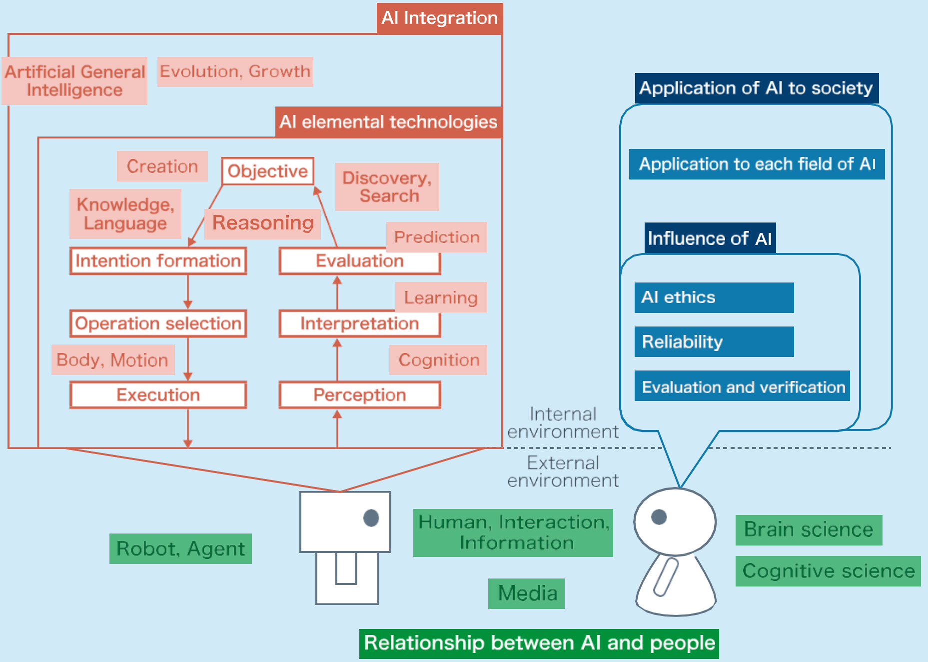
Flow of intelligence activity

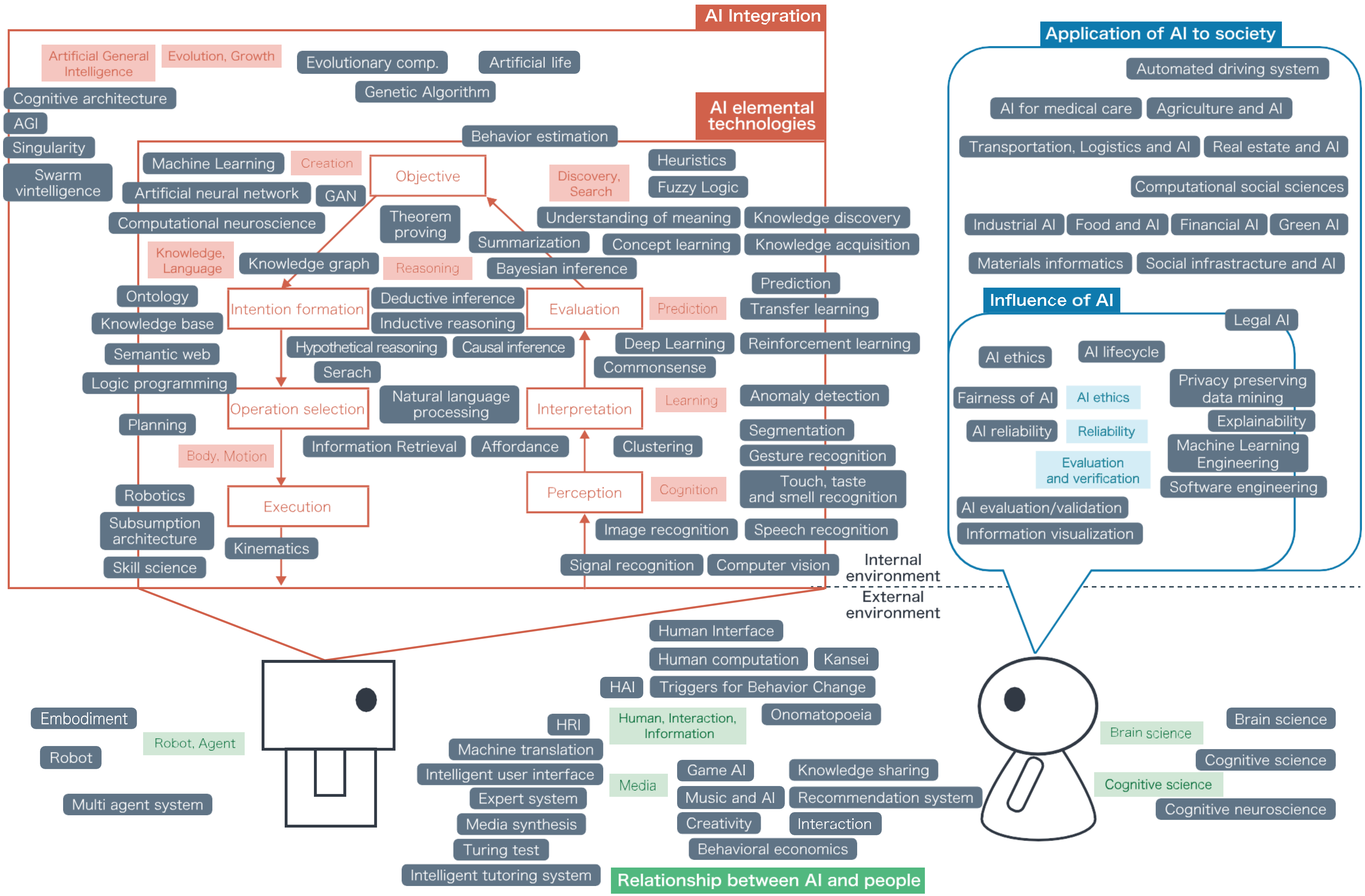
AI research sees human activity as a flow consisting of a combination of many intellectual activities. There are research fields that correspond to each step in this flow. Humans perceive and interpret the visual image, pay attention to the required information, evaluate the information based on the selected information, form an intention, and decide a series of operation sequences. For example, let's consider a fellow researcher who approaches while holding out his right hand. I recognize the right hand approaching and identify the person as non-Japanese. In addition, his expression is friendly. I remember that there is a custom of shaking hands in foreign countries. I combine the recognition, and construct a series of actions, such as putting out my right hand, smiling while making eye contact, and shaking his hand.

AI also needs to work with humans by communicating with the people around it, and this involves many areas of research. For example, one area of research studies the interaction and dialogue between humans and robots with physical bodies.

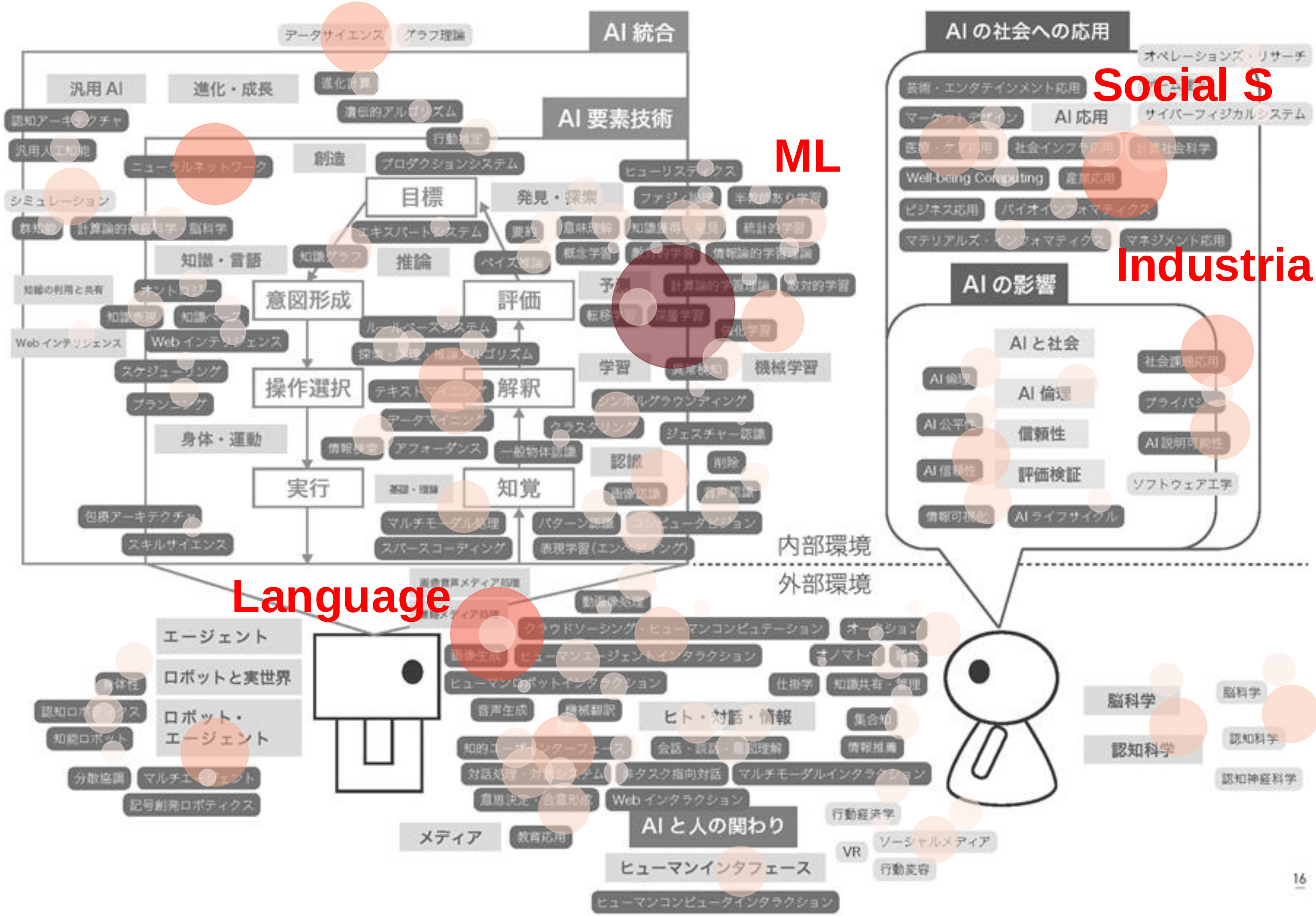
In addition, many new research fields are emerging that examine how humans view AI. Research is also required on the appropriate use of AI, and includes evaluating AI's reliability and operability.

Novices can learn about applications and activities related to their academic fields. For those who are already researching a certain field, the map can show related AI research themes and highlight possible partners for collaboration.



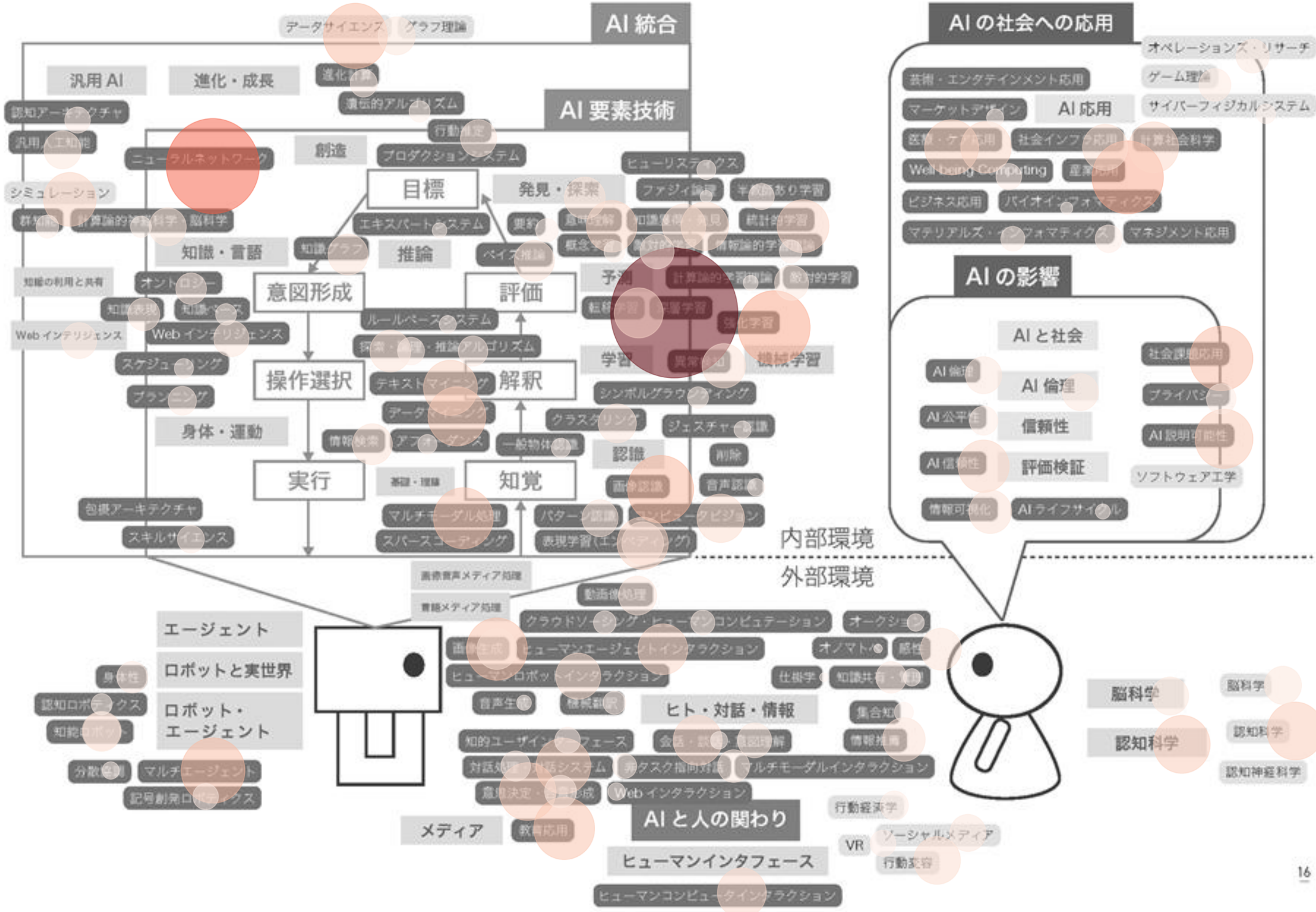

AI Integration
Artificial General Intelligence
Evolution, Growth
Evolutionary comp.
Artificial life
Genetic Algorithm
Cognitive architecture
AGI
Singularity
Swarm vintelligence
AI elemental technologies
Behavior estimation
Machine Learning
Creation
Objective
Discovery, Search
Heuristics
Fuzzy Logic
Artificial neural network
GAN
Theorem proving
Computational neuroscience
Understanding of meaning
Knowledge discovery
Summarization
Concept learning
Knowledge acquisition
Knowledge, Language
Knowledge graph
Reasoning
Bayesian inference
Prediction
Ontology
Intention formation
Deductive inference
Evaluation
Prediction
Transfer learning
Knowledge base
Inductive reasoning
Semantic web
Hypothetical reasoning
Causal inference
Deep Learning
Reinforcement learning
Serach
Commonsense
Logic programming
Operation selection
Natural language processing
Interpretation
Learning
Anomaly detection
Planning
Segmentation
Body, Motion
Information Retrieval
Affordance
Clustering
Gesture recognition
Robotics
Execution
Perception
Cognition
Touch, taste and smell recognition
Subsumption architecture
Image recognition
Speech recognition
Kinematics
Skill science
Signal recognition
Computer vision
Internal environment
External environment
Application of AI to society
Automated driving system
AI for medical care
Agriculture and AI
Transportation, Logistics and AI
Real estate and AI
Computational social sciences
Industrial AI
Food and AI
Financial AI
Green AI
Materials informatics
Social infrastructure and AI
Influence of AI
Legal AI
AI ethics
AI lifecycle
Privacy preserving data mining
Fairness of AI
AI ethics
Explainability
AI reliability
Reliability
Machine Learning Engineering
Evaluation and verification
Software engineering
AI evaluation/validation
Information visualization
Human Interface
Human computation
Kansei
HAI
Triggers for Behavior Change
Embodiment
Robot, Agent
Robot
Multi agent system
HRI
Human, Interaction, Information
Onomatopoeia
Machine translation
Intelligent user interface
Media
Game AI
Knowledge sharing
Expert system
Music and AI
Recommendation system
Media synthesis
Creativity
Interaction
Turing test
Behavioral economics
Intelligent tutoring system
Relationship between AI and people
Brain science
Brain science
Cognitive science
Cognitive science
Cognitive neuroscience

JSAI 2024



AI 統合
AI 要素技術
AI の社会への応用
AI の影響
AI と人の関わり
Social S
ML
Industrial
Language
データサイエンス
グラフ理論
汎用 AI
進化・成長
進化計算
遺伝的アルゴリズム
認知アーキテクチャ
汎用人工知能
ニューラルネットワーク
創造
プロダクションシステム
行動推定
シミュレーション
目標
発見・探索
計算論的神経科学
脳科学
知識・言語
推論
エキスパートシステム
知識グラフ
ベイズ推論
オントロジー
意図形成
評価
予測
計算論的学習理論
敵対的学習
転移学習
深層学習
強化学習
Web インテリジェンス
ルールベースシステム
探索・推論アルゴリズム
スケジューリング
プランニング
操作選択
テキストマイニング
解釈
学習
機械学習
データマイニング
シンボルグラウンディング
身体・運動
アフォーダンス
情報検索
クラスタリング
ジェスチャー認識
一般物体認識
認識
実行
知覚
画像認識
音声認識
包摂アーキテクチャ
スキルサイエンス
マルチモーダル処理
スパースコーディング
パターン認識
コンピュータビジョン
表現学習
ヒューリスティクス
統計的学習
情報論的学習理論
内部環境
外部環境
オペレーションズ・リサーチ
芸術・エンタテインメント応用
マーケットデザイン
AI 応用
サイバーフィジカルシステム
社会インフラ応用
計算社会科学
Well-being Computing
産業応用
ビジネス応用
バイオインフォマティクス
マテリアルズ・インフォマティクス
マネジメント応用
AI と社会
AI 倫理
社会課題応用
プライバシー
AI 公平性
信頼性
AI 説明可能性
AI 信頼性
評価検証
ソフトウェア工学
情報可視化
AI ライフサイクル
画像音声メディア処理
動画像処理
クラウドソーシング・ヒューマンコンピュテーション
オークション
エージェント
ロボットと実世界
ロボット・エージェント
身体性
認知ロボティクス
知能ロボット
分散協調
マルチエージェント
記号創発ロボティクス
画像生成
ヒューマンエージェントインタラクション
ヒューマンロボットインタラクション
音声生成
機械翻訳
ヒト・対話・情報
知識共有・管理
仕掛学
集合知
情報推薦
知的ユーザーインターフェース
会話・談話・意図理解
対話処理・対話システム
タスク指向対話
マルチモーダルインタラクション
意思決定・合意形成
Web インタラクション
行動経済学
メディア
教育応用
ヒューマンインタフェース
VR
ソーシャルメディア
行動変容
ヒューマンコンピュータインタラクション
脳科学
認知科学
認知神経科学

JSAI 2023



AI 統合
データサイエンス
グラフ理論
汎用 AI
進化・成長
進化計算
遺伝的アルゴリズム
認知アーキテクチャ
汎用人工知能
AI 要素技術
行動推定
創造
ニューラルネットワーク
プロダクションシステム
目標
発見・探索
ヒューリスティクス
ファジィ論理
半教師あり学習
シミュレーション
エキスパートシステム
要約
知識獲得・発見
統計的学習
群知能
計算論的神経科学
脳科学
概念学習
情報論的学習理論
知識・言語
知識グラフ
推論
ベイズ推論
予測
計算論的学習理論
敵対的学習
知識の利用と共有
オントロジー
意図形成
評価
転移学習
強化学習
知識表現
知識ベース
ルールベースシステム
Web インテリジェンス
探索・推論・推論アルゴリズム
学習
異常検知
機械学習
スケジューリング
操作選択
テキストマイニング
解釈
シンボルグラウンディング
プランニング
データマイニング
クラスタリング
ジェスチャー認識
身体・運動
情報検索
アフォーダンス
一般物体認識
認識
削除
実行
基礎・理論
知覚
画像認識
音声認識
包摂アーキテクチャ
マルチモーダル処理
パターン認識
コンピュータビジョン
スキルサイエンス
スパースコーディング
表現学習
内部環境
外部環境
画像音声メディア処理
言語メディア処理
動画像処理
クラウドソーシング・ヒューマンコンピュテーション
オークション
エージェント
画像生成
ヒューマンエージェントインタラクション
オノマトペ
感性
ロボットと実世界
ヒューマンロボットインタラクション
仕掛学
知識共有・管理
身体性
認知ロボティクス
音声生成
機械翻訳
ヒト・対話・情報
集合知
ロボット・エージェント
知能ロボット
知的ユーザインタフェース
情報推薦
対話処理
対話システム
マルチモーダルインタラクション
分散協調
マルチエージェント
意思決定・合意形成
Web インタラクション
記号創発ロボティクス
AI と人の関わり
行動経済学
メディア
教育応用
VR
ソーシャルメディア
行動変容
ヒューマンインタフェース
ヒューマンコンピュータインタラクション
AI の社会への応用
オペレーションズ・リサーチ
芸術・エンタテインメント応用
ゲーム理論
マーケットデザイン
AI 応用
サイバーフィジカルシステム
医療・ケア応用
社会インフラ応用
計算社会科学
Well-being Computing
産業応用
ビジネス応用
バイオインフォマティクス
マテリアルズ・インフォマティクス
マネジメント応用
AI の影響
AI と社会
社会課題応用
AI 倫理
AI 公平性
信頼性
プライバシー
AI 説明可能性
AI 信頼性
評価検証
ソフトウェア工学
情報可視化
AI ライフサイクル
脳科学
認知科学
認知神経科学

Automated Diagnosis (2)

- Diagnosis from symptoms and laboratory examinations (Classical ML)
  - Almost solved in 2000’s.
  - The methodologies apply to E-commerce in Google and Amazon.
    - Collaborative filtering, Personalization
  - However, diagnosis from images and/or waveform : still difficult in 2000’s.

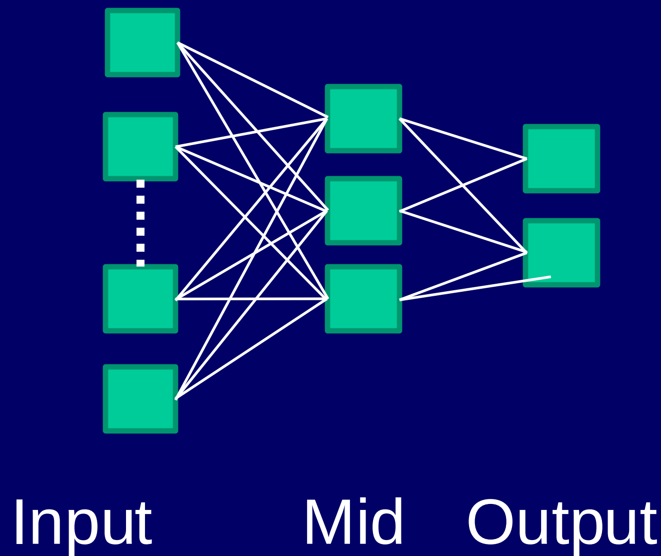
Automated Diagnosis (3)

- Deep Learning improves the precision of image diagnosis in 2010’s.

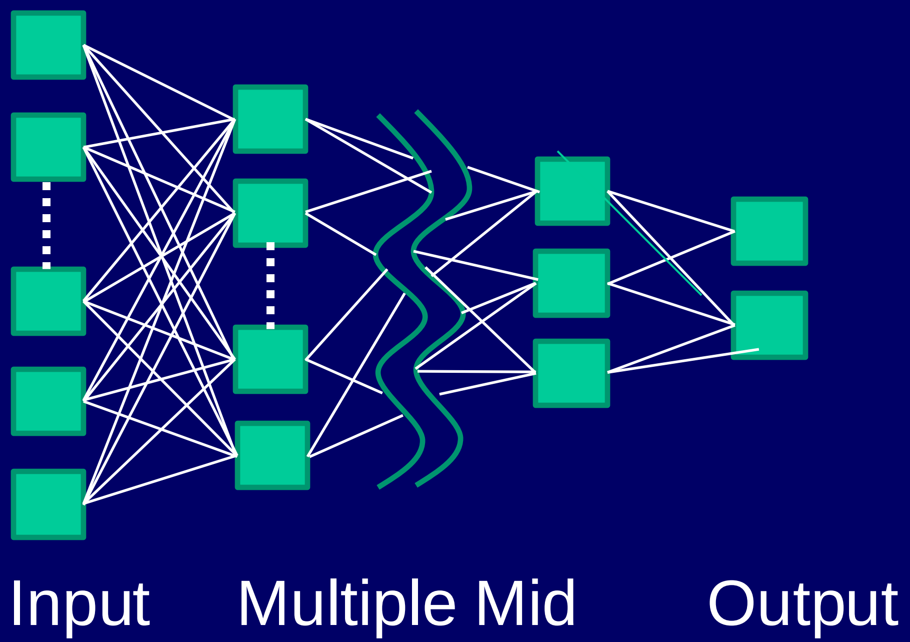
- Method:
  - Developed filtering technologies are integrated into one packages: reuse.

    - Hint: vision recognition model
    - Transfer learning

Neural Network

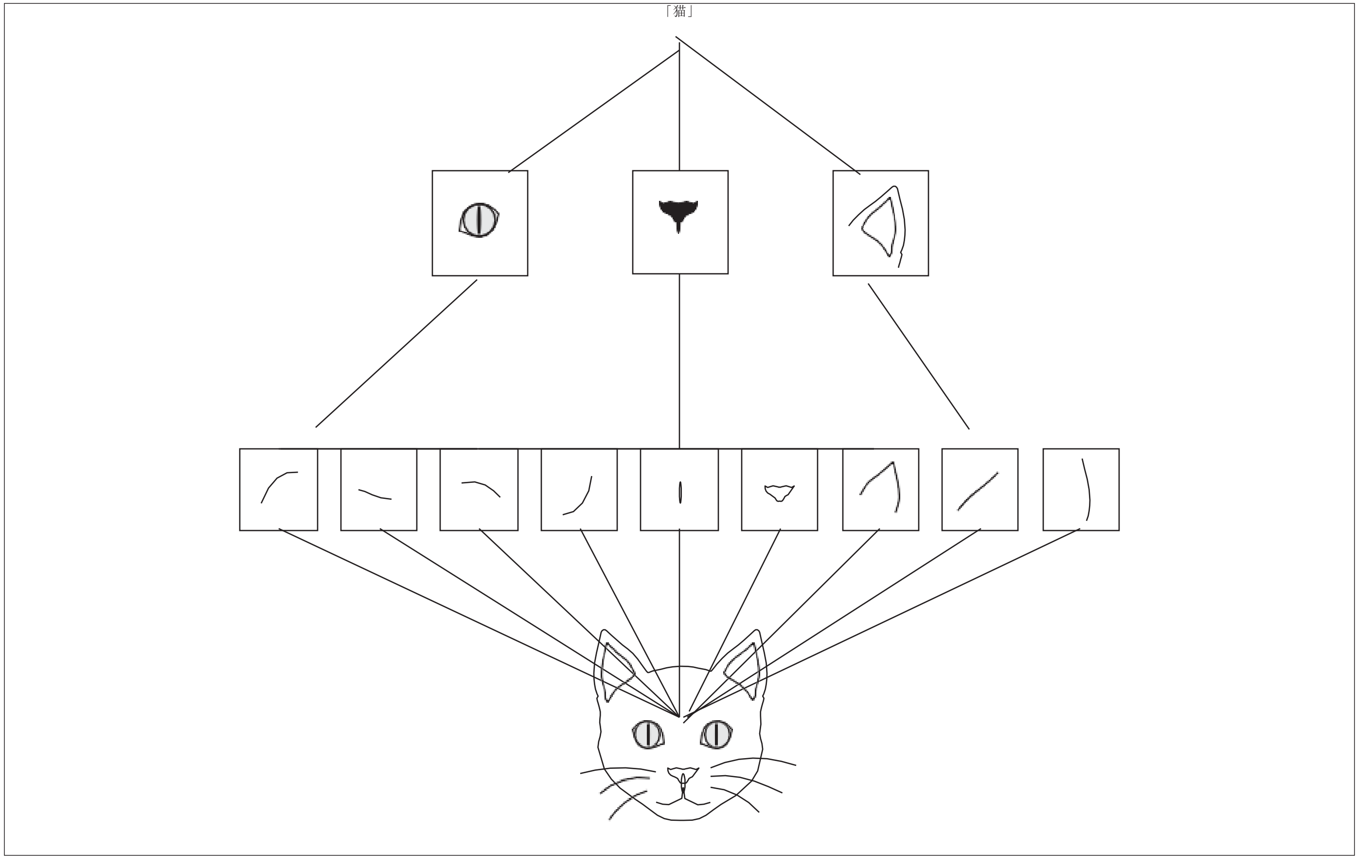


Backpropagation



Deep Neural Network

Vision Recognition Model



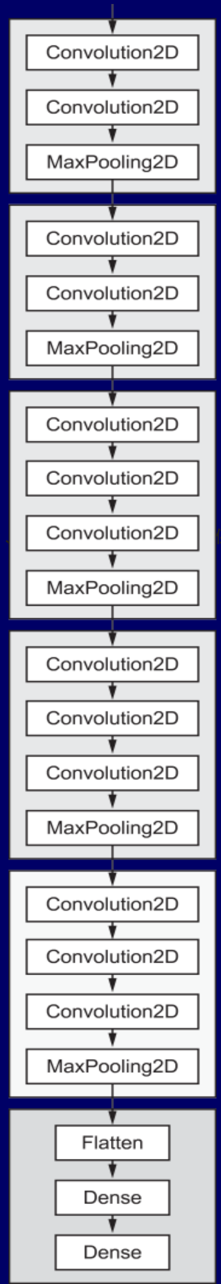
(Deep Learning with R and Keras)

Deep Learning

- CNN
  - Image recognition: higher than previous methods (70%=> 90%)
- High Precision: voice Recognition, wave analysis
  - Image, Sound, ECG, …..
- Pre-Training + Fine tuning
  - available for future use

- For logical analysis, the performance is not as good as conventional methods.

Fine Tuning: Reuse of Past Learning Results



Convolution Layer1：Freezed

Conv. Layer 2：Freezed

Conv. Layer 3: Freezed

Conv. Layer 4: Freezed

Conv. Layer 5: Fine Tuning

Output Layer: Fine Tuning

(Deep Learning with R and Kerase)

Generative AI

- Generation
  - Generate new knowledge from existing data and knowledge

- Present generative AI:
  - Generation model: images,….
    - Diffusion Model
  - Transfer model: machine translation
    - Transformer, GPT, ChatGPT

Machine Translation (1)

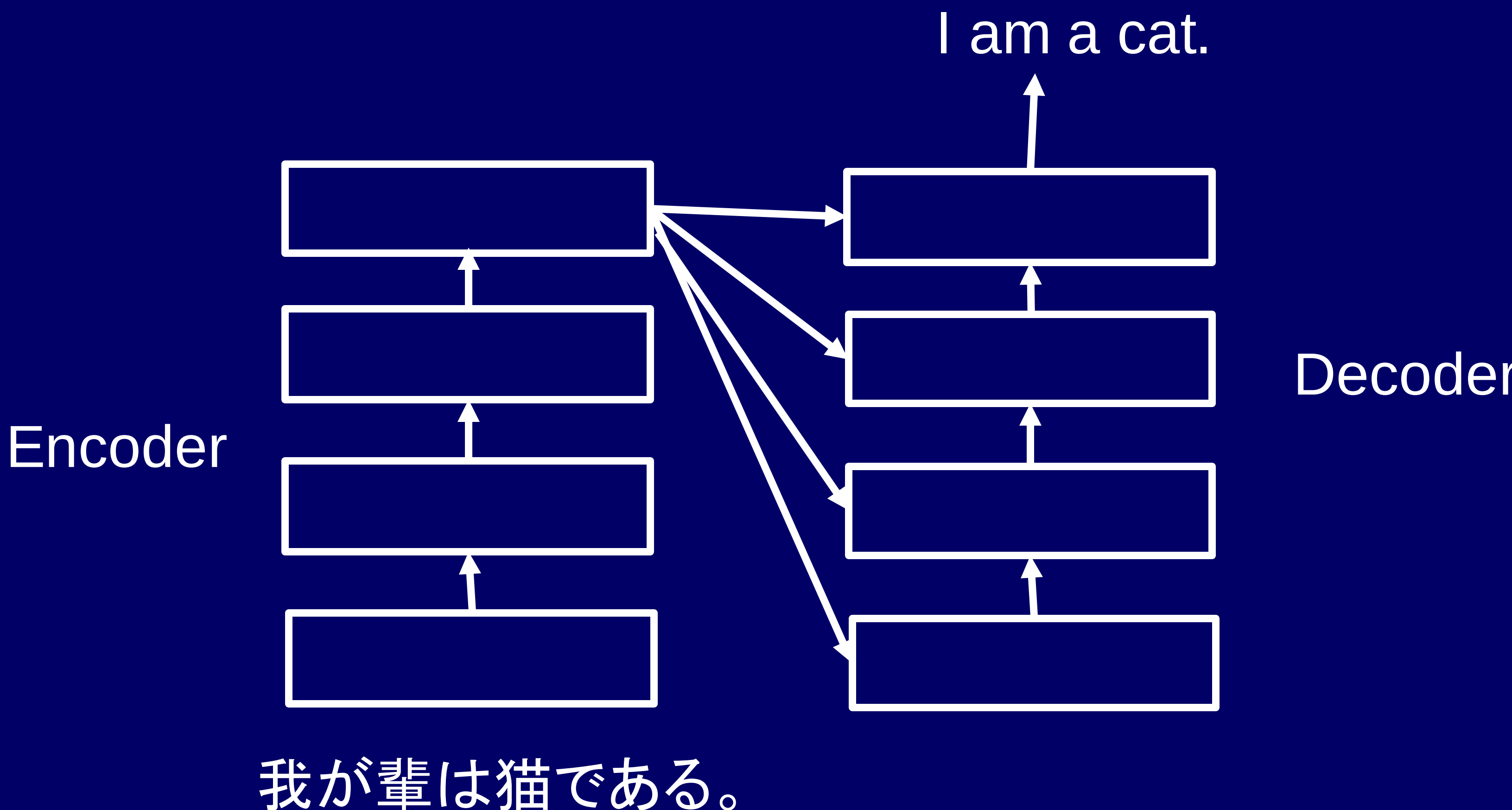
- Japanese → English
- 我が輩は猫である。→ I am a cat.
- Process:
  - Syntactic analysis/ morphological analysis (segmentation of sentences)
  - => Search for Corresponding Words
  - => Generation of English Sentence

Machine Translation (2)

- 1990's : Probabilistic model for syntax Analysis (Hidden Markov model)
- 2000's: Linguistic databases (Corpus)
  Vector representation
- 2010's: Deep learning
  - First, RNN/LSTM (sequential)
  - Transformer / Encoder-Decoder
  - BERT, GPT-2
- Points: words are represented by numeric vectors.

Machine Translation (3)

- Transformer (2017): Google



Machine Translation (4)

BERT (2018):

- Pre-training + Fine Tuning
- Pre-training: general linguistic rules
- firstly exceeds Human score at the benchmark.

GPT-2 (2019):

- Only Decoder-type
- Number of Parameter

GPT-3 (2020):

- #Parameter: 170billion→ ChatGPT

Learned general linguistic rules: Fundamental model of language ?

Interactive AI

- One Goal is to make a Chatbot

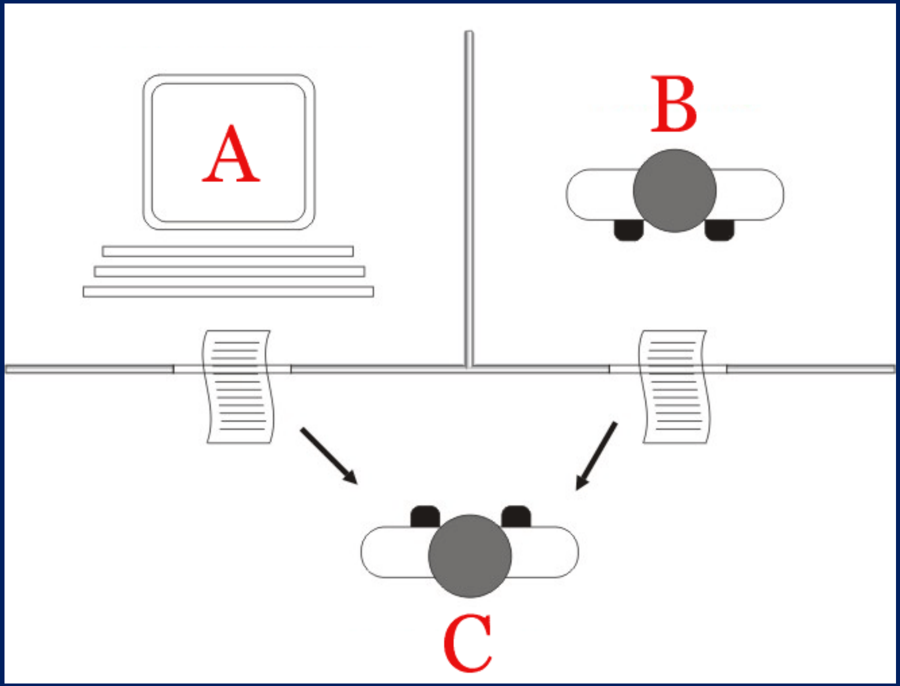
- 1950’s: ELIZA: only keywords bounced back
- ChatGPT (2022)
  - Almost subjects are covered
  - It seems that ChatGTP will pass the Turing test

Interactive AI

- Turing Test (1950)
  - Proposed by Alan Turing

- The Turing Test, proposed by the British mathematician and computer scientist Alan Turing in 1950, is a measure of a machine's ability to exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human. In the test, a human evaluator interacts with both a machine and a human through a computer interface without knowing which is which. If the evaluator cannot reliably tell the machine from the human based on their responses, the machine is considered to have passed the test and demonstrated human-like intelligence. The Turing Test is a fundamental concept in the field of artificial intelligence and explores the possibility of machines thinking and understanding like humans.

Turing Test



The "standard interpretation" of the Turing test, in which player C, the interrogator, is given the task of trying to determine which player – A or B – is a computer and which is a human. The interrogator is limited to using the responses to written questions to make the determination.

From Wikipedia

Interactive AI

- One Goal is to make a Chatbot

- 1950’s: ELIZA: only keywords bounced back
- ChatGPT (2022)
  - Almost subjects are covered
  - It seems that ChatGTP will pass the Turing test

Chat GPT

- It has been expected that learning from large scale data will make a good interactive AI system
  - ⇒ but troublesome
- Open AI uses GPT to learn from 10,000 books and internet data and develops an interactive system (ChatGPT).
- The results are more than expected.
  - Performance of translation: significantly improved.
  - Answers all varieties of questions.
  - Acquires “World model” based on linguistic data => Large Linguistic Model (LLM)

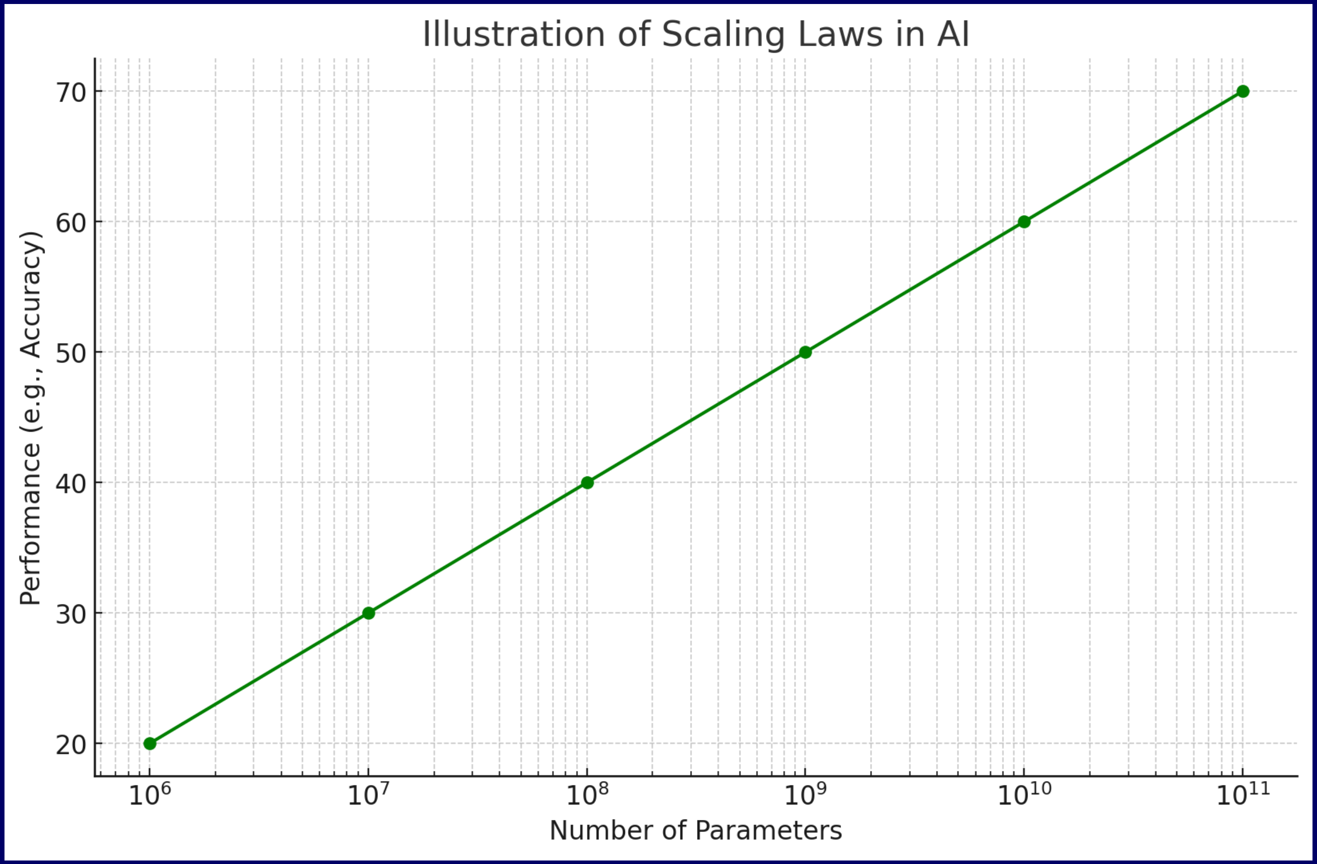
Interestingly,

- Chat GPT almost passes the Turing Test
- Chat GPT shows the validity of Tractatus Logico-Philosophicus, by Ludwig Wittgenstein.
- Sentence (segmentation) => Word Usage => Logical Form=> Sentence Generation

3.326 In order to recognize the symbol in the sign we must consider the signicant use.
3.327 The sign determines a logical form only together with its logical syntactic application.
3.328 If a sign is not necessary then it is meaningless. That is the meaning of Occam's razor.
(If everything in the symbolism works as though a sign had meaning, then it has meaning.)
3.33 In logical syntax the meaning of a sign ought never to play a rôle; it must admit of being established without mention being thereby made of the meaning of a sign; it ought to presuppose only the description of the expressions.

Chat GPT (2)

- Scaling Law: exponential growth of parameters will gain the performance exponentially.



Large Linguistic Model (LLM)

- All the linguistic info are represented by numerical vectors.
- LLM stores almost all the information of vectors.
  - Relations between words.

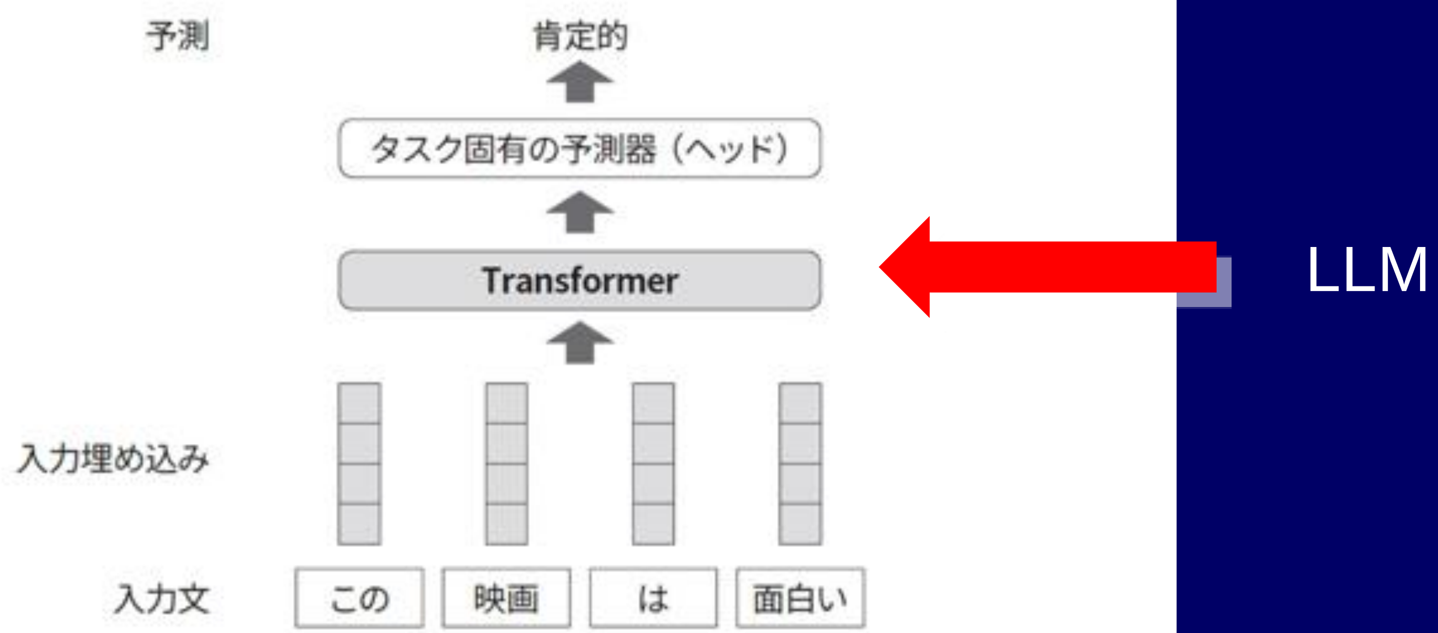


図 1.4: Transformer を使った大規模言語モデルをファインチューニングして感情分析を解く例。灰色の部分が事前学習されている。

Before LLM

- We need to make a program that learns knowledge about relations of words.

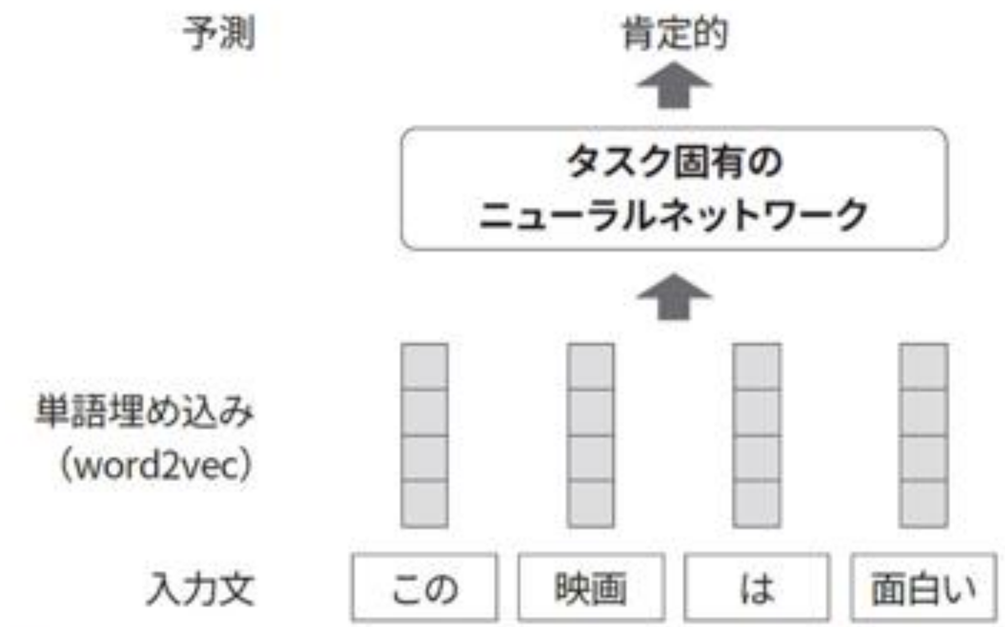


図 1.3: 単語埋め込みを使って感情分析を解く例。タスク固有のニューラルネットワークの入力として、word2vec で訓練した単語埋め込み（灰色の部分）を用いる。

What we learn from LLM

Various representation can learn and store as templates of vectors.

"It can store and learn various expressions like templates. It can also learn which words to fill in the templates → meaning that it can also learn dependency structures and more."

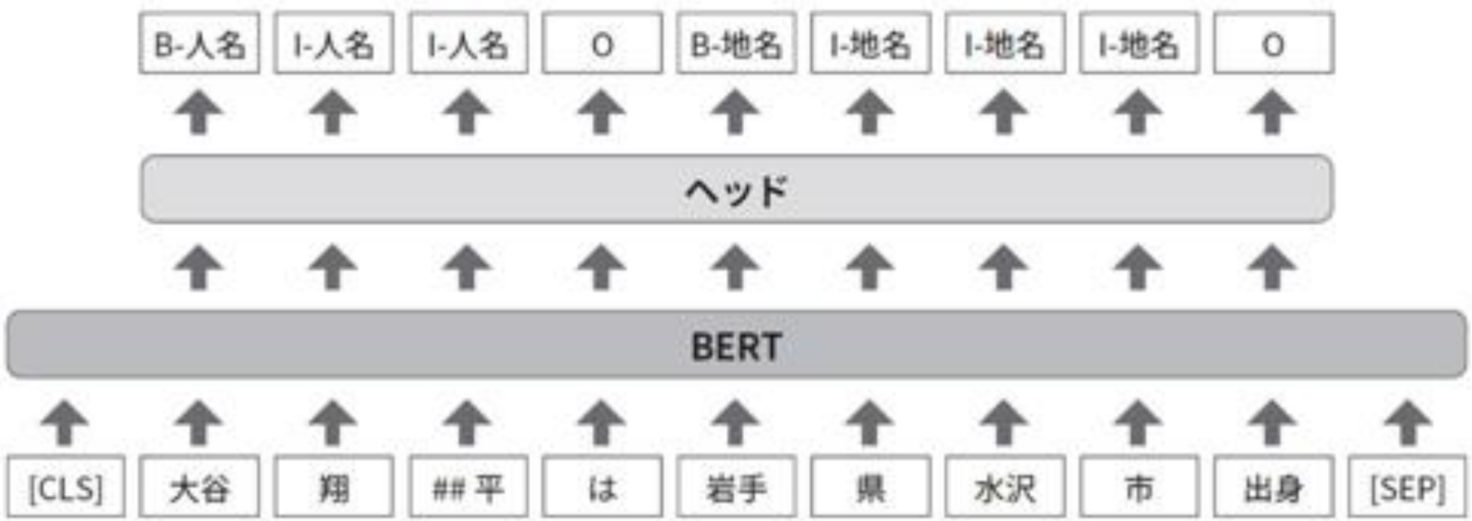


図 3.7: BERT を使った固有表現認識のファインチューニング

Four Clasical Problems in AI

- Intelligent game:
  - Almost achieved.

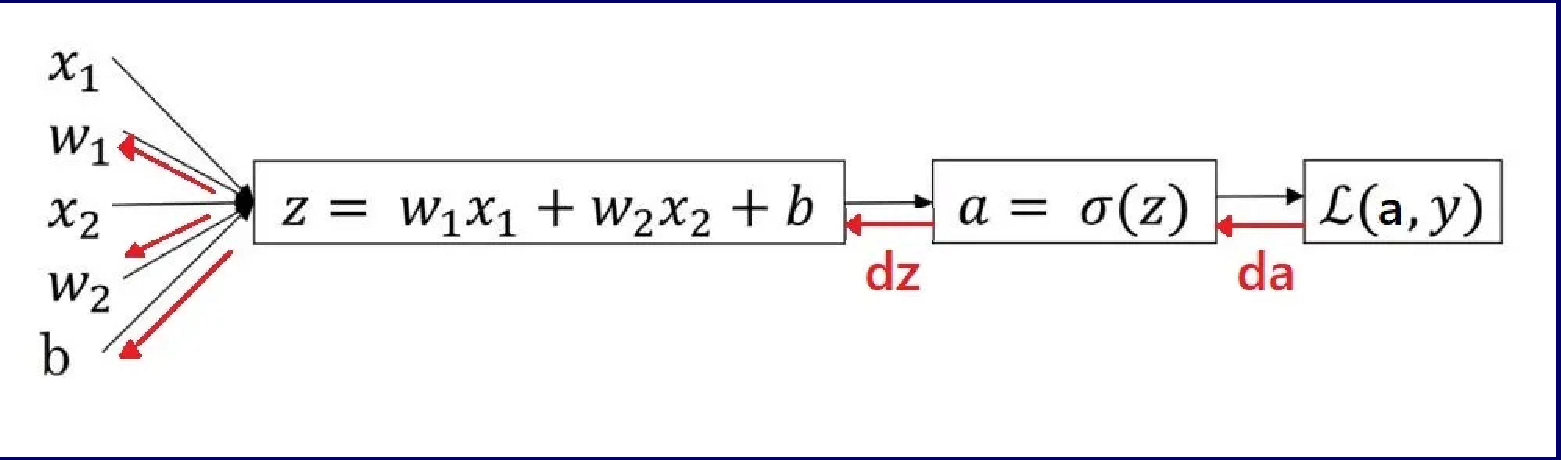
- Problem dolving
  - Image and waveform also improved
- Machine translation
  - Improved
- Turing test (Interactive AI):
  - Almost Pass?

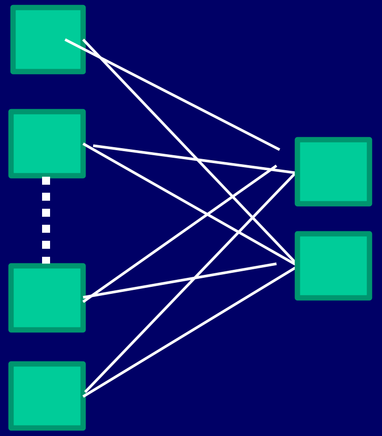
Future ?

Future for Data Processing

- AI for Data Preprocessing
  - Calculation of propensity score
    - From logistic to deep learner
  - Transformation of data
    - From Text to Data (by LLM?)
  - Multi-modal analysis
    - From Image/Audio to Text
  - (Fine-Tuning: Transfer Learning)

Logistic Regression as Perceptron





- Logistic Regression:
  Classical Perceptron

- Logistic regression used for
- calculation of propensity score
- => It may be strengthen by deep learning method

LLM for Data Transfomration

- LLM can be used for extraction of data from texts.
- LLM can learn the transformation of data format.

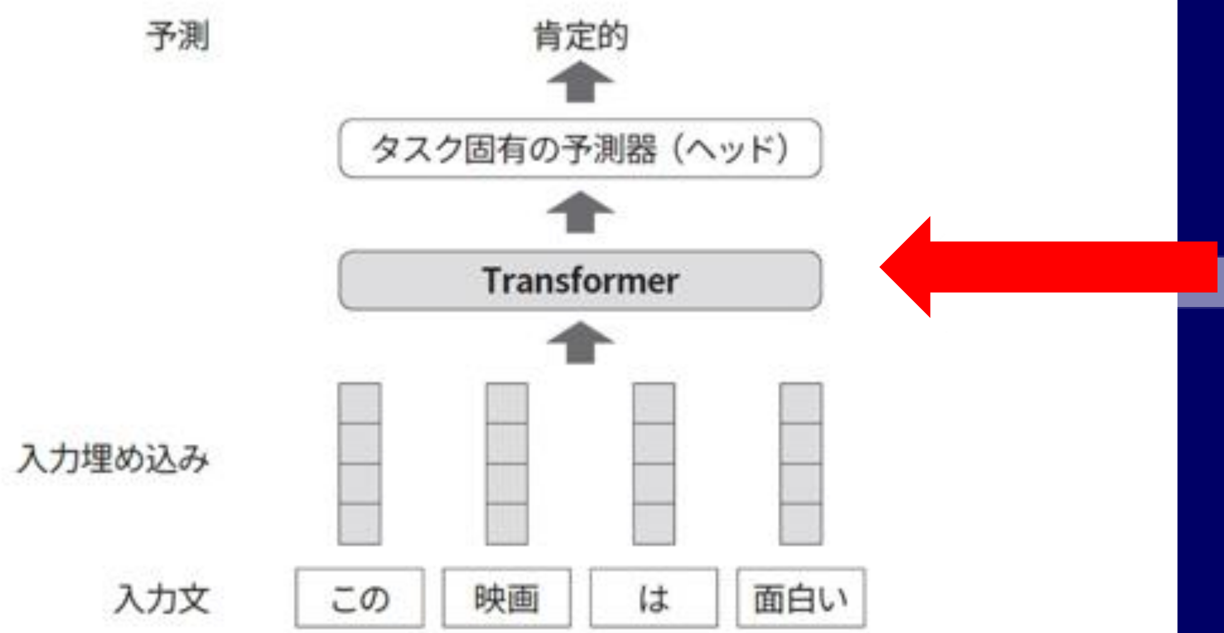


図 1.4: Transformer を使った大規模言語モデルをファインチューニングして感情分析を解く例。灰色の部分が事前学習されている。

Toward (Multi-modal)Diagnosis

nature medicine

Review Article | Published: 07 January 2019

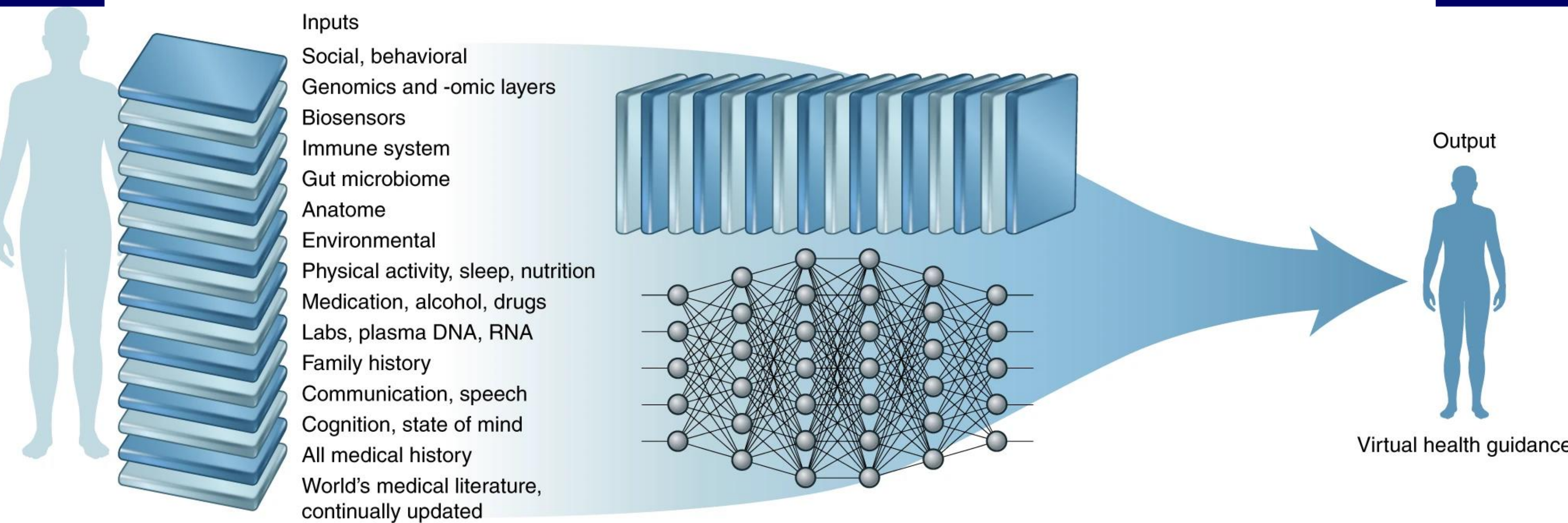
High-performance medicine: the convergence of human and artificial intelligence

Eric J. Topol

Nature Medicine **25**, 44–56 (2019) | Cite this article

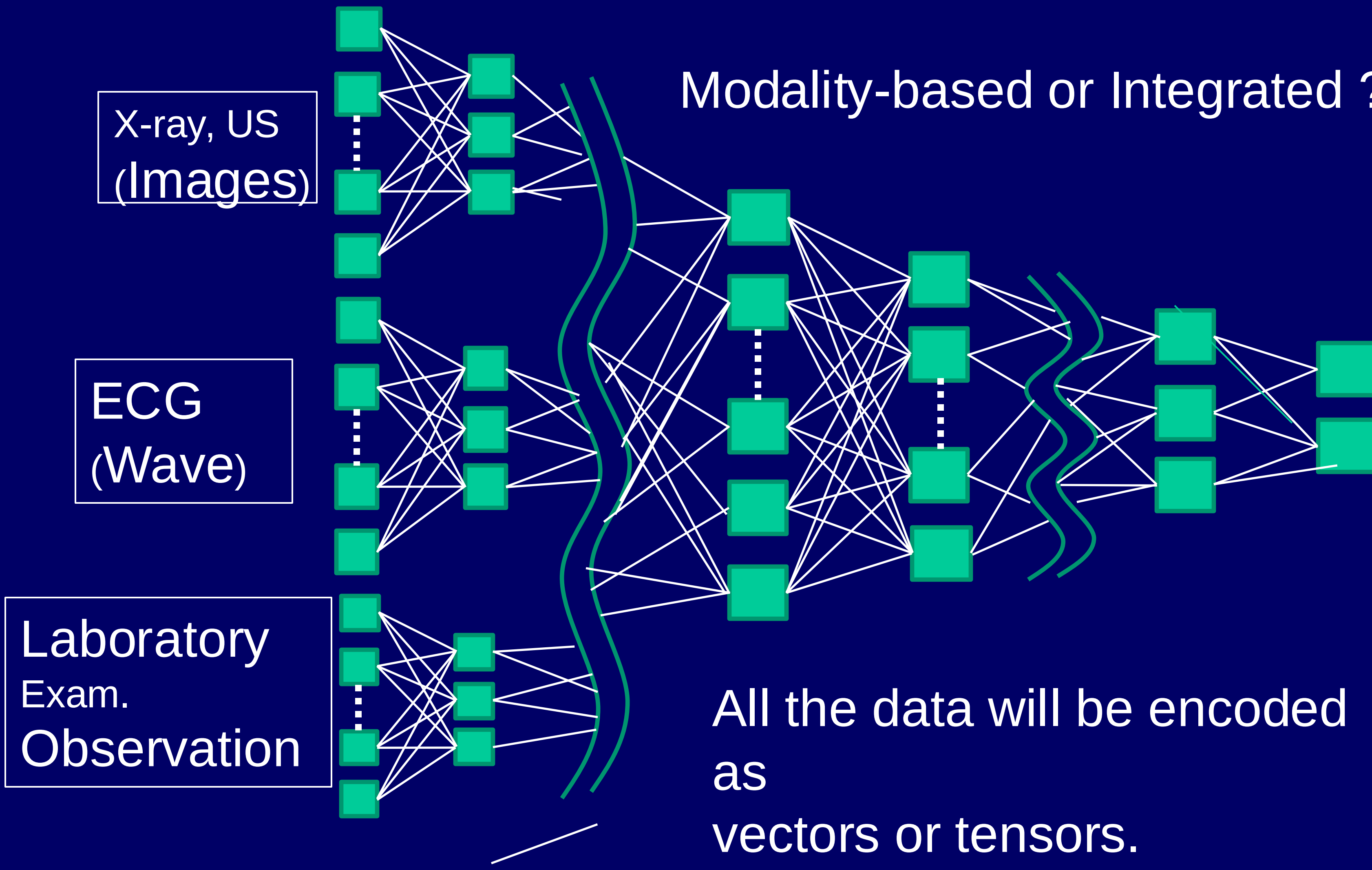
Associated content

Innovations In



Multi-Modal Deep Learning (Proposal)

Modality-based or Integrated ?



All the data will be encoded as
vectors or tensors.

Image + Text

- Input
  - X-Ray images
  - Radiological Reports (Text)
- Learning medical visual representation
- Image Encoder + Text Encoder
  - Interaction between encoders
    - Cross-Attention
- Target
  - Medical Image Classification
  - Medical Object Detection
  - Medical Semantic Segmentation

Image + Text

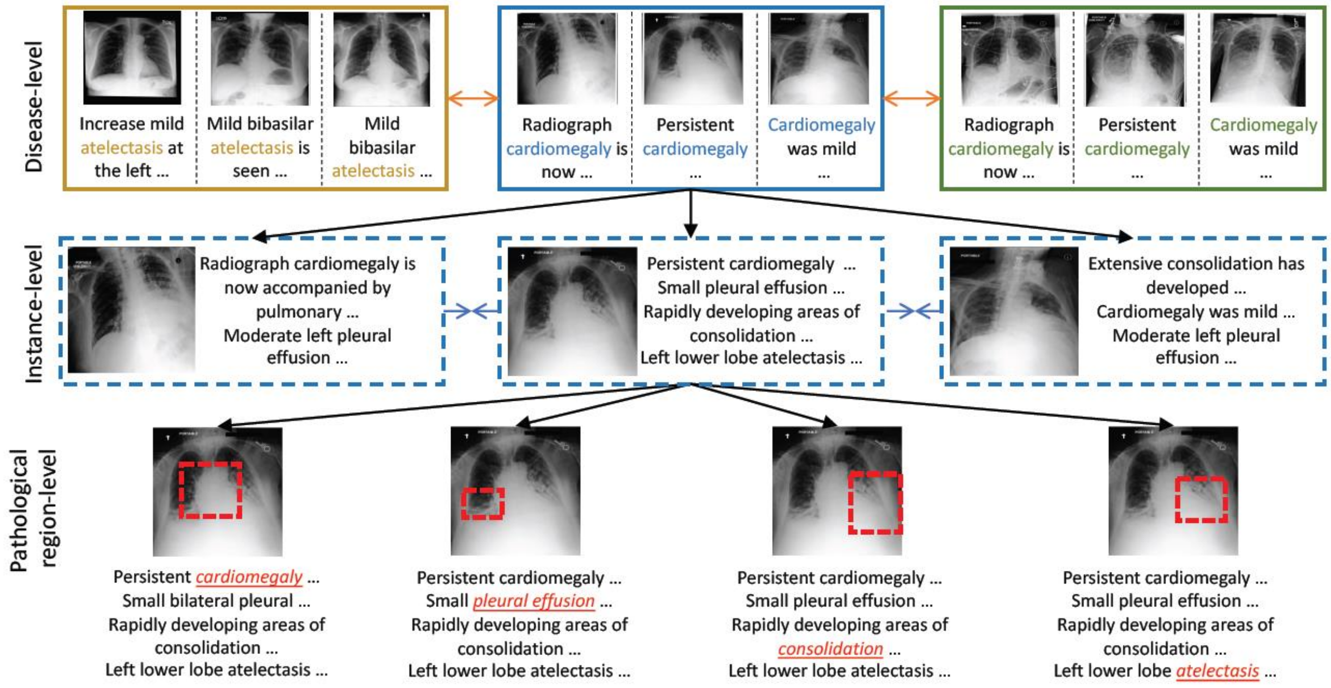


Figure 1: The multi-granularity (disease-level, instance-level, and pathological region-level) semantic correspondences across medical images and radiology reports.

Multi-Granularity Cross-modal Alignment for Generalized Medical Visual Representation Learning

Fuying Wang, Yuyin Zhou, Shujun Wang, Varut Vardhanabhuti, Lequan Yu

arXiv: 2210.06044v1 12 Oct 2022 (NeuroIPS 2022)

Image + Text

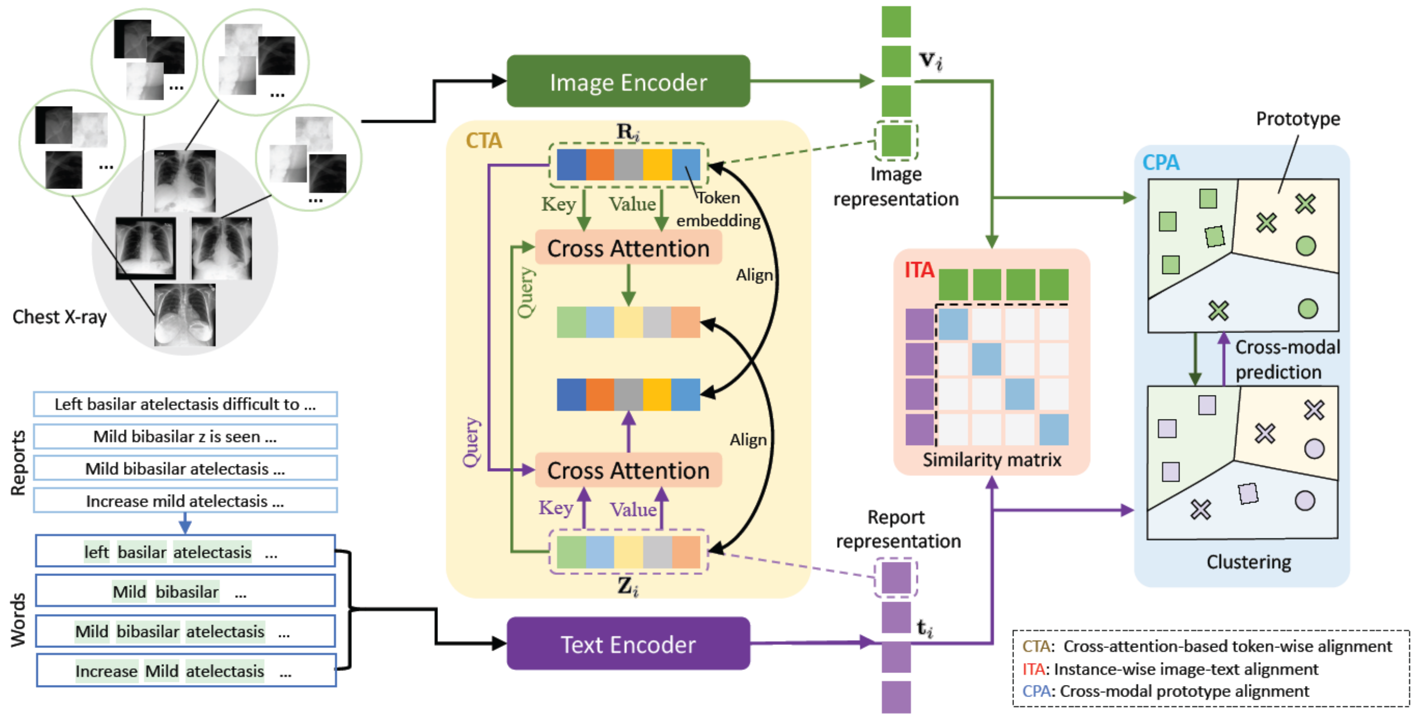


Figure 2: Illustration of our proposed multi-granularity cross-modal alignment framework. CTA, ITA, and CPA represent token-wise alignment, instance-wise alignment, and prototype (disease)-level alignment respectively. The green arrow represents information flow of visual features, while the purple arrow represents information flow of textural features.

Multi-Granularity Cross-modal Alignment for Generalized Medical Visual Representation Learning

Fuying Wang, Yuyin Zhou, Shujun Wang, Varut Vardhanabhuti, Lequan Yu

arXiv: 2210.06044v1 12 Oct 2022 (NeuroIPS 2022)

Summary

- Artificial Intelligence
  - Four classical problems
- Machine Learning
  - has empowered AI systems
- Generative AI
  - ChatGPT (LLM)
- Future: AI for Data Preprocessing