

G検定対策講座 第4回

知識表現と エキスパートシステム

第二次AIブームの中心となった技術



今日学ぶこと

1

知識表現の基本

意味ネットワーク・オントロジー・セマンティックウェブ

2

エキスパートシステムとは

知識ベース + 推論エンジン

3

代表的なシステム

ELIZA・DENDRAL・MYCIN

4

大規模な知識システム

Cyc・ワトソン・東口ボくん

5

データマイニング

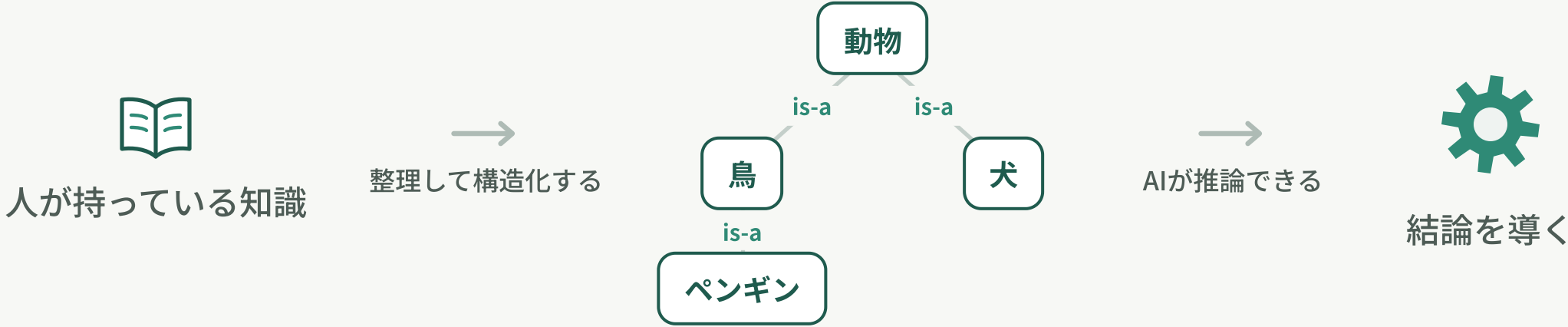
データから知識を掘り出す

AIに「知識」を 与えるとは？

コンピュータが扱える形で、概念と関係をあらわすこと

知識表現とは

コンピュータが扱える形で、概念や関係をあらわす技術



POINT AIが考えるには、まず知識を整理して構造化する必要がある。これが第二次AIブームの中核



1968年 キーリアン

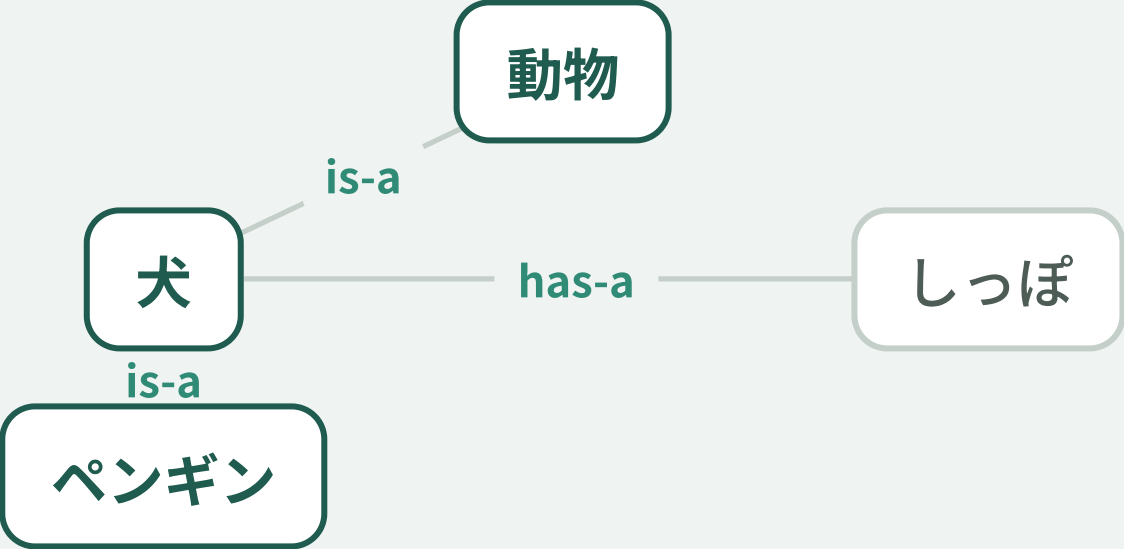
意味ネットワーク

概念を点（ノード）、関係を線（リンク）であらわすグラフ構造



キーリアンが提唱した
知識表現の方法

具体例 よく出る3つの関係



is-a / 種類（犬は動物の一種）

has-a
/ 所有（犬はしっぽを持つ）

part-of / 部分
（しっぽは犬の一部）

POINT この3つを、次のスライドから詳しく見ていく



関係その1

is-a 関係（種類）

A は B の一種であるという関係

上位の性質を下位が引き継ぐ
これを継承と呼ぶ

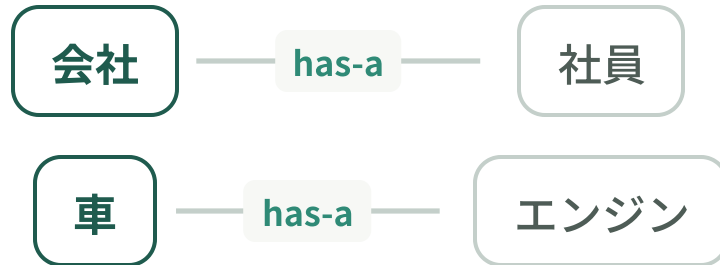
具体例 継承のしくみと、その落とし穴

鳥が卵を産むなら
ペンギンも産む、と推論できる（継承）でも「鳥は飛べる」は継承できない。
すべての性質が継承できるとは限らない

POINT この落とし穴（例外の扱い）は、試験でもよく問われる

has-a 関係 と part-of 関係

似ているが、切り離せるかどうかが違う

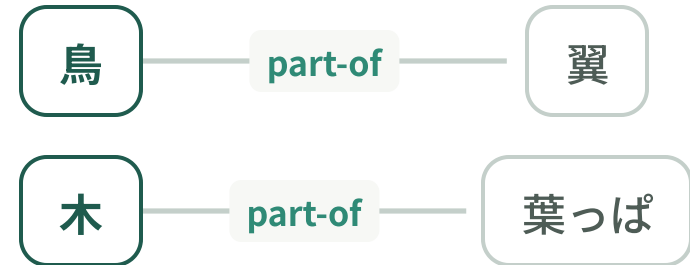


has-a (所有)

A は B を持っている

会社がつぶれても社員は別の会社へ移れる

≠



part-of (部分)

B は A の一部である

鳥がいなくなれば翼もなくなる

知識の設計図



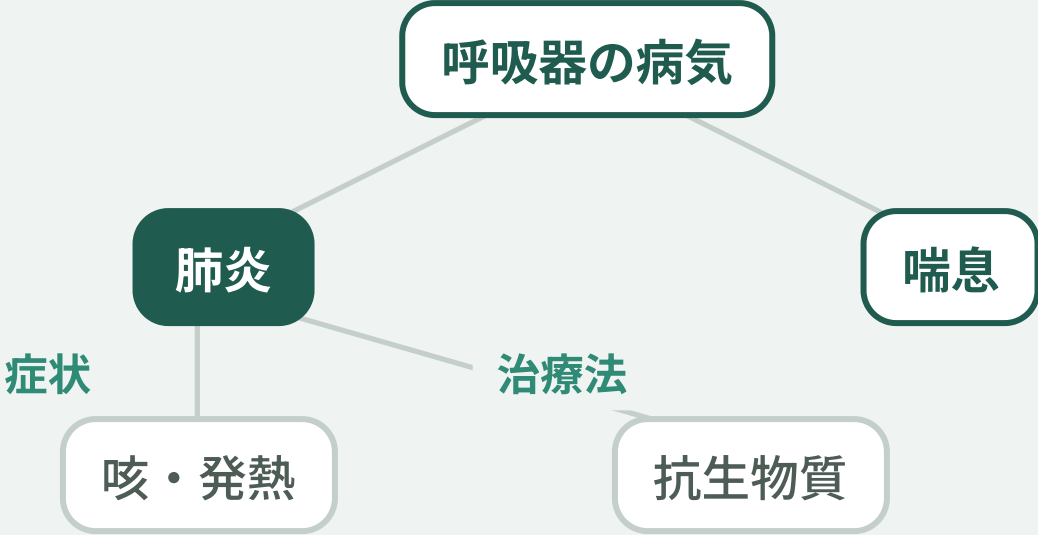
オントロジー

ある分野の概念の体系を、
階層という形で書きあらわしたもの



もとは哲学の存在論。
AIでは知識ベースを作る枠組み

具体例 医療のオントロジー



概念どうしの関係を
体系的に書いておく

同じオントロジーを使えば
別のシステム同士でも意味が通じる

POINT 生命科学のジーン・オントロジーをはじめ、医療やネット通販でも実用化されている

構想



セマンティック・ウェブ

機械が意味を理解できるウェブに
しよう、という構想



ウェブの父
ティム・バーナーズ＝リーが提唱

具体例

いまのウェブとの違い

今 ふつうのウェブ

人間が読むためのもの

機械には意味のわからない記号の並び

案 セマンティック・ウェブ

ページに意味のタグを付けておく

「30分の和食レシピ」を自動で集められる

関係を書く RDF

オントロジーを書く OWL

POINT

この考え方は、いまの検索エンジンの土台にも受け継がれている

第二次AIブーム



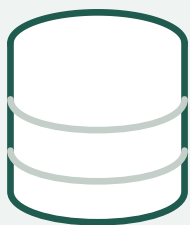
エキスパートシステム

専門家の知識をコンピュータに入れて、
推論させるAI



質問に答えると診断が返ってくる
相談・応答の形をとる

具体例 2つの部品でできている



知識ベース
If-Thenルールや事実



推論エンジン
ルールから結論
を導く



使われ方

質問に答えると
診断が返ってくる
限定された分野では
人間の専門家に匹敵した

POINT 知識ベースと推論エンジンの組み合わせ。第二次AIブームの主役



1966年

ELIZA（イライザ）

エキスパートシステム以前の、
対話の先駆け



マサチューセッツ工科大学の
ジョセフ・ワイゼンバウム作

具体例 ロジャース派の精神科医をまねる

利用者 悲しい

ELIZA なぜ悲しいのですか

仕組み 言葉を入れ替えて返すだけ



機械的な受け答えなのに
なんだか人間らしく感じる

これがイライザ効果という言葉の語源

POINT 中身は単純なパターンマッチング。本物のAIではない



1965年～

DENDRAL (デンドラル)

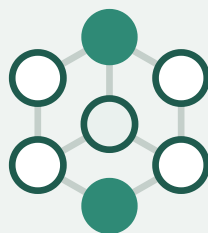
世界初のエキスパートシステムとされる、化学の構造解析AI



スタンフォード大学
ファイゲンバウムらが開発

具体例 質量分析データから、構造を推定する

質量分析の
データ



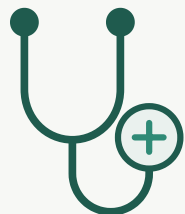
「これはステロイドの構造ですね」
と答えを出せる

化学者の推論を
数百のルールに変換した

実験の測定結果

有機化合物の
構造を推定

POINT 専門家の頭のなかの手順を、ルールとして書き出した最初の成功例



1972年～

MYCIN（マイシン）

感染症の診断を行うエキスパートシステム



スタンフォード大学。
ショートリッフェらが開発

具体例 数百個のIf-Thenルールで診断する

入力 発熱・せき・血液検査の結果

推論 細菌性肺炎の疑いあり

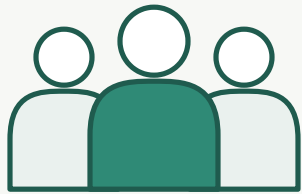
出力 ペニシリンの投与を提案



確信度を数値で扱った
「どのくらい確からしいか」

誤診の責任という問題があり
実用化はされなかった

POINT 診断の精度は専門医に匹敵。医療AI倫理の議論の出発点にもなった



知識獲得

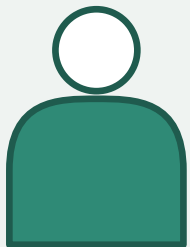
インタビューシステム

専門家に聞き取りをして、**ルール**を取り出す手法



この作業をする人を
知識エンジニアと呼ぶ

具体例 聞き取りから、ルールへ



専門家

→
聞き取る

質問 肺炎を疑うのはどんな時か

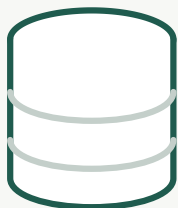
回答 発熱とせきが続き、白血球が多い時

ルール もし発熱＋せき＋白血球 ならば 肺炎



これを全部くり返す
地道すぎる作業

POINT ベテラン医師の「なんとなくおかしい」は**ルール**にできない。これが第二次AI冬の原因



1984年～

Cyc（サイク）プロジェクト

AIに人間の常識をすべて与えようとする壮大な試み



ダグ・レナトが開始
いまだ40年以上続いている

具体例 入力しているのは、こういう知識

人は呼吸する

水は濡れる

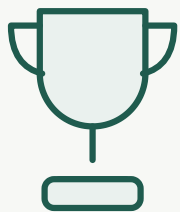
亡くなった人には再会できない



数百万の事実と
ルール

当たり前すぎて
誰も書かない知識を
ひたすら人手で入力し続ける

POINT 常識をどう扱うかという問題に、真正面から挑み続けている



2011年

ワトソン (Watson)

IBMが開発した質問応答システム



クイズ番組ジェパディ！で
人間のチャンピオンに勝利

具体例

仕組みと、その後

自然言語の質問

ふつうの言葉で聞く



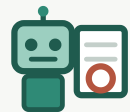
大量の文書から
作った知識ベース



候補を探して
確からしさを採点

POINT

その後医療や金融へ応用され、ワトソン・ヘルスなどのサービスも展開された



2011～2016年

東ロボくん

東京大学の合格を目指した、日本のAIプロジェクト



国立情報学研究所
新井教授が主導

対比 できたこと / 越えられなかった壁

○ できた

センター試験でMARCHレベルに到達
明治・青山学院・立教・中央・法政

× 越えられず

東大合格は断念
読解力の壁を越えられなかった

POINT AIは意味を理解しない。のちの大規模言語モデル時代の読解研究の出発点にもなった



知識の掘り出し

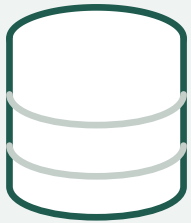
データマイニング

大量のデータから、
有用なパターンや知識を掘り出す技術



ウェブ上のデータなら
ウェブマイニング

具体例 有名な例：ビールと紙おむつ



購買履歴の
大量データ



パターンを探す



見つけたパターン

**ビールと紙おむつが
一緒に売れている**

人間には思いつかない組み合わせ

POINT 1990年代以降、人手で入力する時代からデータから自動で見つける時代へ。その転換の象徴

