JDLA Deep Learning for General 2017　勉強会資料

# 目次

[目次 1](¥l)

[0. はじめに 2](¥l)

[1. ディープラーニングの概要 3](¥l)

[1.1 ニューラルネットワークとディープラーニング，既存のニューラルネットワークにおける問題，ディープラーニングのアプローチ 3](¥l)

[1.2 CPUとGPU 4](¥l)

[1.3 ディープラーニングにおけるデータ量 4](¥l)

[2. ディープラーニングの手法 5](¥l)

[2.1 活性化関数 5](¥l)

[2.2 学習率の最適化 6](¥l)

[2.3 CNN 7](¥l)

[2.4 RNN 8](¥l)

[2.5 深層強化学習 9](¥l)

[2.6 深層生成モデル 10](¥l)

[3. ディープラーニングの研究分野 11](¥l)

[3.1 画像認識 11](¥l)

[3.2 自然言語処理 12](¥l)

[3.3 音声処理 14](¥l)

[3.4 ロボティクス（強化学習） 15](¥l)

[3.4 マルチモーダル 15](¥l)

[4. ディープラーニングの応用に向けて 17](¥l)

[4.1 産業への応用 17](¥l)

[4.2 法律 19](¥l)

[4.3 倫理 24](¥l)

[文献目録 25](¥l)

# はじめに

検定・資格概要，試験範囲，例題，推薦図書から，以下のような観点で問題が出題される可能性が高いと推測しました．

1. モデルや手法の基本的な説明（数理的知識のない人に説明できる表現を用いたもの）
2. 各モデル・手法が用いられている理由
3. ディープラーニングを用いる際に注意すべきこと
4. 各トピックの今までの動向（特に最新動向が出題される可能性が高い）
5. ディープラーニングでできること，事業への応用事例

この資料では，上記の観点から，出題範囲のトピックに対する知識をまとめました．また，例題・試験範囲と推薦図書を見比べた結果，推薦図書に載っていない内容が出題される可能性があると判断したので，私たちの判断で推薦図書にない内容も一部載せています．下記の章は，「試験範囲だが推薦図書にない部分」で，ネット情報や自分たちの知見を元にまとめた部分になります．

1.2.1 CPUとGPUの違い

1.3 ディープラーニングにおけるデータ量

3.4 マルチモーダル

上記以外の範囲に関しては，参考図書の内容を要約したものを載せています

# ディープラーニングの概要

## 1.1 ニューラルネットワークとディープラーニング，既存のニューラルネットワークにおける問題，ディープラーニングのアプローチ

### 1.1.1 基本的な言葉の説明

**ニューラルネットワーク**：生物の神経回路網を模倣した数理モデル．

**ディープラーニング（深層学習）**：深い層を重ねることで学習精度を上げるように工夫したニューラルネットを用いる機械学習技術

**誤差逆伝播法**：サンプルに対するネットワークの誤差（目標出力と実際の出力との差）を，入力層から逆に伝播させ，各層の重みの勾配を計算するという方法．

**勾配消失問題**：誤差逆伝播法の重みの勾配の計算の際，入力層から離れた深い層に進むに連れ，勾配が急速に小さくなったり，あるいは大きくなって発散してしまう現象．この問題が多層ニューラルネットの学習を困難なものにしていた．

**事前学習**：目的とするネットワークの学習の前に，パラメータのよい初期値を得るための教師なし学習．自己符号化器による手法と制約ボルツマンマシンによる手法がある．

(1) pp.1-5

### 1.1.2 ニューラルネットワーク，ディープラーニング研究の今までの動向

研究のブームは現在を含めて３回である．２回目のブームは，誤差逆伝播法の発明によるものだが，以下２つの理由からブームが下火になってしまった．

①誤差逆伝播法は，多層になるとうまくいかない．

勾配消失問題にもとづく過学習が問題

②学習パラメータ（層数，ユニット数等）の性能への寄与がよくわからない．

その後，事前学習や新たな活性化関数の研究によって勾配消失問題がある程度解決されたことで深い層の学習が可能になり３回目（現在）のブームが起きた．

(1) pp.1-5

### 1.1.3 ディープラーニングを用いる際に注意すべきこと

・過学習

・訓練データの量が足りないと過学習を起こしやすい

解決策として手持ちのデータを加工し「水増し」する（**データ拡張**）ことがある．

(1)p.35

## 1.2 CPUとGPU

### 1.2.1 CPUとGPUの違い

**CPU** ：PC上の様々な情報を処理するために使われる．コア数が少なく（2~9個程度），並列計算に向いていない．

**GPU** ：数千のコアが搭載されていて，並列計算に特化している．

### 1.2.2 最新動向

・GPUを利用することが一般的である．

・スーパーコンピュータ技術を利用した様々な形態の**並列化**が試みられている．中でも，プロセッサや計算ノードをまたぐ分散並列化には，次のようなものがある．

**データ並列**：複数の計算機が同一のDNNのパラメータを持ち，異なるデータについて並列に計算を行う手法

**モデル並列**：複数の計算機が同一のデータを分割して持ち，異なるモデルの箇所について並列に計算を行う手法．

・今後の展望として，中期的には，消費電力低減やメモリ容量に対する要求から， **FPGA**の利用や**専用チップ**の開発の推進が求められている．

**FPGA**：ディープラーニングなどの処理をハードウェアに実装する．CPUやGPUに比べて消費電力を抑えられるという点で優位．

・さらに長期的には**量子コンピュータ**等の発展も念頭に置いておく必要がある．

(2) pp.105, 115-116．

## 1.3 ディープラーニングにおけるデータ量

ディープラーニングには大量のデータが必要（例えば，手書き数字の認識を安定させるのに数万データが必要）．少ないデータ数での学習を上手く行うために，以下のような手法がある．

・**データ拡張**：手持ちのデータを加工し「水増し」する手法．

・**転移学習**：学習済みモデルを用いて、「目標のデータ」を学習させる手法（あらかじめ「目標のデータ」と性質の近いデータを使って学習させておく．）

# ディープラーニングの手法

## 2.1 活性化関数

### 2.1.1 活性化関数とは

ニューラルネットワークにおいて，線形変換をした後に適用する非線形関数.

### 2.2.2 代表的な活性化関数

**・ロジスティック関数（ロジスティックシグモイド関数**）：

活性化関数として古くから使われてきた．現在では代わりにランプ関数が使われることが多い．

**・双曲線正接関数：**

ロジスティック関数と同様に，現在では代わりにランプ関数が使われる．

**・正規化線形関数（ReLu,ランプ関数）：**

近年で最もよく使われる関数．値域に上限がないため，勾配消失問題が起きにくい．

**・マックスアウト関数：**

複数のユニットが出力した値について，最大値を取り出す関数．CNNでは使われているが（**最大プーリング層**のこと），それ以外では一般的に使われない．

**・恒等写像：**

活性化関数は基本的には非線形関数だが，部分的に恒等写像を使うことがある．(回帰問題の場合は，出力層に使われる．)

**・ソフトマックス関数：**

クラス分類を目的とするネットワークの場合，出力層にこの関数が使われる．(出力値を確率として解釈できるようになる．)

**・LReL (leaky rectified linear)関数：**

2013年に考案された関数．ランプ関数と比べて良いかは不明確．

(1)p.10

## 2.2 学習率の最適化

### パラメータの最適化手法について

・確率的勾配降下法(SGD，逐次的勾配降下法)：

訓練データ1個ずつに対して誤差関数(損失関数)を求め，パラメータの更新を行っていく手法．バッチ学習と比べて局所的な解に陥りにくい．

・バッチ学習（エポック学習）：

全データの誤差関数の和を求め，まとめてパラメータの更新を行う．局所的な解に陥りやすい．

・ミニバッチ学習：

複数個のデータの誤差関数の和を求めて，パラメータの更新を行う手法(SGDとバッチ学習の中間のような手法)．SGDと比べて計算効率が良い．

上記3つの中で,ミニバッチ学習が最もよく使われる．

### 学習率(学習係数)とは

・パラメータの更新量の大きさを決める係数のこと．

・　この式ののことを指す．（：重み，：損失関数）

### 学習率の決め方

・AdaGradという手法により，学習率を自動で変化させることができる．

・学習の途中で，手動で変えていくことも多い．

### 最新動向

・勾配降下法の収束性能を向上させる(局所解に陥らないようにする)ための手法として**モーメンタムSGD**という手法がある．

・「ミニバッチ学習でAdaGradを用いる」か「ミニバッチ学習でモーメンタムSGD」が主流．（後者の場合，学習率を手動で変えていく必要あり．）

## 2.3 CNN

### 2.3.1 CNNとは

・生物の脳の視覚野の仕組みを元に作られた順伝播型ネットワークモデル．

・主に画像認識に応用されている．

・**畳み込み層**と**プーリング層**という特殊な層を使っている．

**畳み込み層**：画像データに対して，一定範囲ごとに特徴量を抽出し，特徴の位置を表す画像（特徴マップ）を出力する．

**プーリング層**：畳み込み層から出力された特徴マップに対して，一定範囲ごとに代表値（最大値など）を取り出していくことで，圧縮（低次元化）された画像を出力する．①計算量の軽減と②特徴の多少の傾きや位置のずれを許容できるという2つのメリットがある．

### 2.3.2 CNNの利点

全結合でなく，層同士を部分的に結合させることで，局所的な特徴を抽出することができる．

### 2.3.3 よく使われているモデルや手法

・基本的には，畳み込みとプーリングを繰り返し行う構造が使われる．

・出力層に近い層には**全結合層**を採用するのが一般的．

・自然画像（写真など）を扱う場合，入力画像の全体的な明るさやコントラストの違いを吸収するために，局所コントラスト正規化（local contrast normalization）という手法が使われる.

## 2.4 RNN

### 2.4.1 RNN(リカレントニューラルネット,再帰型ニューラルネット)とは

・内部に（有向）閉路を持つニューラルネットの総称．

・サンプルごとに長さの異なるデータや，系列内の並び（文脈など）に意味のあるデータに対して有効なモデル．

・音声や言語，動画像といった系列データに対して適用されている.

### 2.4.2 RNNの利点

内部に閉路を持つことで，過去に入力した情報を保持することができ，データの並び順の意味を抽出できる．

### 2.4.3 RNNの逆伝播計算

・RNNの学習の際に誤差の微分を計算する方法は，RTRL法とBPTT法の２種類がある．

・RNNは，長い系列データ（長い文章など）に対して，勾配消失問題を起こしやすい（長期間の過去情報を保持できない）．その解決策として，**LSTM（Long Short-Term Memory）**が考案された．

### 2.4.4 LSTM(Long Short-Term Memory)について

・RNNの中で現在，最も良く使われるモデル.

・LSTMは，RNNに忘却ゲートという仕組みを取り入れたモデルで，勾配消失問題が起きにくくなり，長期的な過去情報も保持できるようになった．

### 2.4.5 RNNの応用

RNNは，入力データの長さと同じ長さの出力を返す（入力した文章と同じ数の単語を出力する）．入力の長さと異なる長さの出力をしたい場合，以下の手法が提案されている．

・**隠れマルコフモデル**：RNNと組み合わせることも可能．

・**コネクショニスト時系列分類法（CTC）**：RNNの出力の解釈を変更したもの．

## 2.5 深層強化学習

### 2.5.1 強化学習とは

・エージェント（学習の主体）が，環境から現在の状態を認識し,試行錯誤を繰り返しながら，報酬が最大になるような行動を学習する．

・学習データを事前に用意する必要がない．

・強化学習にディープラーニング（主にCNN）を組み込むことで，様々なボードゲームでプロを超えるアルゴリズムが作られるようになった（アルファ碁がきっかけで，強化学習における深層学習の応用がより発展した）．

### 2.5.2 代表的な手法

**・Q-学習　・TD法　・Sarsa法**

**・深層Q学習**：Q学習にCNNを組み込んだモデルのこと．行動価値関数（行動に対する報酬の期待値）を計算する際に，関数近似としてCNNを用いている．Atari2600の半数以上のゲームでプロの75％以上の成績を達成し，既知の手法を大幅に上回る成果を残した．

Atari2600：米国アタリ社が開発した家庭用ゲーム機．スペースインベーダーやパックマンが有名．

**・モンテカルロ木探索：**アルファ碁は，この手法に加えてCNNによる教師あり学習（プロの棋譜の学習）を組み合わせており，2016年にイ・セドル九段に勝利した．

### 2.5.3 応用事例

・ゲーム

・車，ロボットの制御

・ファイナンスへの応用

・広告配信の最適化

## 2.6 深層生成モデル

### 2.6.1 生成モデルとは

・**生成モデル**：

・訓練データを学習し，それらのデータと似たような新しいデータを生成するモデルのこと．教師なし学習．

・高次元のデータから特徴抽出して，データを低次元に落とし込んだ後，その特徴を元にデータを生成する（主成分分析に近い）．

・画像データに対して応用されることが多い．

### 2.6.2 代表的な深層生成モデル

・**制約ボルツマンマシン（RBM）**：訓練データの生成分布を推定するモデル．

・**自己符号化器（Auto Encoder）**：入力データと出力データが直接近くなるように学習を行うモデル．

・**VAE（Variational Auto Encoder）**：自己符号化器の1つ．データ生成時にノイズを加える．

・**GAN（Generative Adversarial Network）**：生成器(generator)と識別器(discriminator)という２つのネットワークを用いた生成モデル．他の手法と比べて高画質な画像を生成できるが，学習が不安定．

### 2.6.3 深層生成モデルの応用

特定のラベルのデータが極端に少ない時や，ラベルのデータが大量にある時に，生成モデルを用いることでデータ数を増やすことがある．医療関係（新薬の探索など）で活用されている.

# 3. ディープラーニングの研究分野

## 3.1 画像認識

ディープラーニングは画像認識分野の中でも，**①クラス分類，②物体検出，③物体セグメンテーション，④画像キャプション生成，⑤画像生成**に応用されている．以下では，それぞれの代表的手法と動向について述べる．

### 3.1.1 クラス分類

・代表的手法：**ResNet**

一般にCNNは層を深くすると学習が困難になるが，深い層でも学習が上手くいくモデルとして，ResNetが提案され，1000層程度の深さでも学習が可能となった．ResNet は，ILSVRC2015においてトップの成績を残している．

・動向

ResNetの登場以降，現在まで大きな発展はしていない．

### 3.1.2 物体検出

**物体検出**：画像内の物体を取り囲むボックスを推定するタスク．

・代表的な手法：**R-CNN**(Regions with CNN)

CNNと線形サポートベクターマシンを組み合わせたモデルである．

・動向

現在は，学習や出力時の計算を高速化した**Faster R-CNN**が提案されている．

### 3.1.3 物体セグメンテーション

**物体セグメンテーション**：対象物と背景を境界まで詳細に切り分けるタスク．

・代表的な手法：**FCN** (Fully Convolution Network)

FCNでは，同じ物体が隣接している場合に，上手く切り分けるのが困難．

・動向

隣接した物体も上手く切り分ける手法として，MNC (Multi-task Network Cascade)などが提案されている．

### 3.1.4 画像キャプション生成

**画像キャプション生成**：画像に対して，自然言語で記述された説明文などを生成すること．(例：「赤い服を着た女性が街中で電話をしている」など)

・代表的な手法

画像をCNNに入力し，CNNから得られた画像特徴をLSTMに入力する．

### 3.1.5 画像生成

代表的な手法

・**Deep Dream**：2015年にGoogleが開発したシステム．通常の画像を夢に出てくるような画像に変換するシステム．

**・Deep Style**：Googleが開発した入力画像の画風を変換するシステム（例：ゴッホ風にする）．AIが製作する芸術作品としてメディアなどで取り上げられた．

・**GAN（生成敵対ネットワーク）**：高画質な画像を生成できるため，現在，最も利用されている．画像生成器と画像識別器から構成されていて，画像生成器は分類器を騙すような画像を生成し，識別器は生成器から生成された画像と本物の画像とを分類するようにそれぞれ学習する．

## 3.2 自然言語処理

自然言語処理におけるディープラーニングによるブレイクスルーは，**分散表現学習**である．

**分散表現**：単語を高次元の実数ベクトルで表現する技術．

分散表現の最も代表的な枠組みは，Googleのトマス・ミコロフ氏が提案した**Word2Vec**である．Word2Vecは，単語ベクトルの演算により，意味的な演算を行うことができる．

例：「王」-「男」+「女」=「女王」

このほか，自然言語処理の様々な分野において，ディープラーニングが用いられるようになった．以下では，その中でも代表的なものについての動向について述べる．

### 3.2.1 構文解析

**構文解析**：自然言語の文の構造を解析する技術．

・単語列の構文的なまとまり(句構造)を木構造で表す方法と,単語の間の文法関係(依存構造)を木構造で表す方法がある．

・最近では，ニューラルネットワークやRNNなどを適用することで更に精度が向上し，90％以上の解析精度が達成されている．

・構文解析は，構文木を正解データとした学習データ（ツリーバンク）が不可欠で，開発にはコストと時間がかかる．

・現在，多数の言語について同じ基準ツリーバンクを開発する「Universal Dependencies」プロジェクトが注目されている．

### 3.2.2 意味構造解析

**意味構造解析**：構文解析の次のステップ．文が表す意味構造を認識する技術．

・構文器（句構造や依存構造）を入力として，意味構造の各要素を機械学習により推定する方法がある．

・単純な分類問題なので，SVM，パーセプトロン，CRF（Conditional Random Fields），RNNなど，様々なモデルを適用させることができる．

### 3.2.3 感情極性解析

**感情極性解析**：文や文章に対して，ポジティブかネガティブかを解析する技術

・古典的な分類問題・構造予測問題に帰着できるため，SVMやCRFなどの基本的なモデルが用いられてきた．

### 3.2.4 文脈解析

**文脈解析**：文章全体の意味を考慮した自然言語処理．代表例は次の２つである．

照応解析：文章内にある代名詞の指し先を明らかにする解析．限られた場面においては実用段階に達しつつある．

談話構造解析：文間の意味的構造を明らかにする解析(原因と結果など)．談話構造が付与されたデータセットPDTB(Penn Discourse Tree Bank)が2006年に公開されたことを皮切りに，盛んに研究が行なわれている．

### 3.2.5 議論マイニング

**議論マイニング**：ディベート等で複数の発言の関係を，議論関係（支持，反論）に分類する．

・談話構造解析に近いが，より複雑な問題になっている．

・機械学習モデルが用いられているが，性能が十分なものでなく，検討の余地は多い．

## 3.3 音声処理

ディープラーニングは音声処理分野の中でも，**①音声認識，②音素モデル，③雑音・残響抑圧，④音声合成**に応用されている．以下では，それぞれの代表的手法と動向について述べる．

### 3.3.1 音声認識

・音声認識では，音声データから得られる音響特徴量（音声の周波スペクトル）を入力として，それをモデルによって音素状態のカテゴリに分類し，発話内容を予測する．

・音声認識のモデルは，従来は広く混合正規分布に基づく隠れマルコフモデル（GMM-HMM）が主流であった．

・これに対して，2010年頃に多段のネットワークを学習する手法が高い認識性能を挙げたことをきっかけとして，DNNが用いられるようになった．

### 3.3.2 音素モデル

・音素モデルとは，音声認識の訓練データを作成する際に，音声データを音素状態に区分するために用いられるモデル．

・音素モデルでは，HMMが未だ主流．

・近年，LSTMを用いて音響特徴量系列から音素や文字などの系列に直接変換するCTC（Connectionist Temporal Classification）という方式が注目を集めている．

### 3.3.3 雑音・残響抑圧

・実環境下で音声認識を行う場合，雑音と残響の抑圧は重要な課題である．

・従来，雑音・残響抑圧では，線形フィルターが主に用いられてきた．

・近年は，高次の非線形フィルターを表現可能などの理由から， DNNが用いられている．

・更に最近では，雑音・残響抑制のDNNと音素モデルのDNNを連結し，統合的に最適化することも研究されている．

### 3.3.4 音声合成

・音声認識の逆過程である音声合成においてもDNNの導入が活発に行なわれてきた．

・通常のニューラルネットワークを用いた音声合成では，入出力の系列が考慮されず，性能が良くなかった．

・しかし，音響特徴量の系列としての特徴を適切にモデル化するためにRNNを用いることで，高品質の音声合成を実現することができた．代表的なものにGoogleの「**WaveNet**」がある．

## 3.4 ロボティクス（強化学習）

ディープラーニングと強化学習を組み合わせた深層Q学習を用いることで，外部から得た複雑なデータを元に強化学習を行うことが可能となり，深層Q学習を用いた自動制御ロボットの研究が行なわれるようになった．

※　推薦図書にも「ロボティクス×強化学習」に関する内容が上記のみであること，ネット情報がほとんど得られなかったことから，非常に内容が少なくなっております．

## 3.4 マルチモーダル

### 3.4.1 マルチモーダル学習とは

**マルチモーダル学習**：複数の種類の情報を入力として取り入れて，モデルの学習をおこなうこと．単一の情報を学習させるよりも，精度や頑健性が向上すると期待されている．

・それぞれの種類の情報は表現が大きく異なるため，単純に結合することはできない．適切な学習ができるようにするためには，共通するより普遍的な特徴表現を獲得する必要がある．

・ロボットなどは，画像だけではなく，音声やセンサ情報など様々な種類の情報を豊富に得るとこができるので，応用が期待されている．

### 3.4.2 代表的なモデル

DNNは，学習の過程において深い層で普遍的な特徴表現が獲得できる．さらに，DNNが教師なし学習で幅広い成功を収めていることから，複数の種類の情報によるDNNの表現学習が提案されてきた．

**多層の自己符号化器（Auto Encoder，AE）**：  
各種類の情報についてAEを用意し，それらの最も深い層を共有することで，共通の特徴表現を獲得する．様々な問題設定による比較検証によって,単一の情報よりも良い特徴表現を獲得できることが示されている.

**DBM（Deep Boltzmann Machine）：**

生成モデルである制約ボルツマンマシン（RBM）を多層にしたモデル．

### 3.4.3 最新動向

VAE（Variational Auto Encoder）を用いたマルチモーダル学習のモデルとして，**マルチモーダルVAE（MVAE）**が提案されている．

# 4. ディープラーニングの応用に向けて

## 4.1 産業への応用

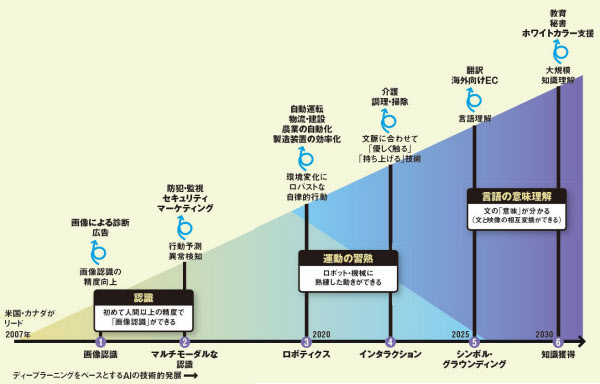
ディープラーニングをベースとした技術は

図 ：ディープラーニングをベースとするAIの技術進展予測

1. **画像や音声等の認識**
2. **運動の習熟**
3. **言葉の意味理解**

の順で発展すると予想されており，導入効果と投資金額のバランスによって段階的に産業への応用が進んでいくと予想される．以下では，認識，運動の習熟，言葉の意味理解の３つの技術の産業応用の動向を示す．

### 4.1.1 認識

**⑴画像認識**

**自動車の走行環境の認識**や**医療分野における画像診断支援**等での産業応用が先行している．

**・走行環境の認識**

「完全自動走行」の実現に向けた取組が進められている．ディープラーニングを中心とした機械学習が自動走行に用いられる理由は，走行環境（特に市街地）は非常に複雑であり，全てのパターンやルールを予め設定して制御することは現実的ではないからである．

**・医用画像の認識**

医用画像の診断支援へのAIの活用による医師不足の解消が期待されている．成功事例の１つであるEnlitic（米国）のがん画像検出システムでは，１人の放射線技師よりも約50％高い精度で，肺の悪性腫瘍を分類することができる．

**・その他**

・防犯カメラから取得した動画像の解析による，都市や施設の安全性を高める取組

・事故画像を解析による損害調査の効率化

**⑵音声・言語認識**

音声認識・対話，機械翻訳，文章生成の領域で産業応用が先行している．

**・音声認識・対話**

大規模コーパスの整備と評価プロジェクトにより音声認識の精度は大幅に向上した．2016年10月にはMicrosoftがCNNとLSTMを組み合わせた認識モデルを開発し，プロの口述記者と同レベルの単語誤り率（Word Error Rate，WER）5.9%を達成．「Siri」や「Google Assistant」のように実用化が進んでいる機能は，ディープラーニングの登場に加え，複数マイクを利用したビームフォーミングや雑音抑制技術の向上等によるものである．

**・機械翻訳**

近年ディープラーニングを用いた機械翻訳の性能が大幅に向上している．代表例である「Google Neural Machine Translation」（GNMT）では，従来のフレーズ単位で翻訳する手法ではなく，ディープラーニングにより文章全体を翻訳単位として捉えることで，英語→スペイン語等の一部のケースにおいて，人間が行う翻訳に近い精度を達成している．

**・文章生成**

画像，テキスト，数値から自然な文章やメッセージ性の高いキーワードを生成する等の技術開発が進んでいる．報道記事の生成やウェブサービスのコンテンツ管理等に利用され始めている．

### 4.1.2 運動の習熟

深層強化学習によりロボットや機械の動作を習熟させる研究開発が始まっている．多くは実証実験・研究開発段階．特に先進的な取組として，以下のようなものがある．

Google：ロボットがオフィスや医療機関等の複雑な現場業務を支援する機能を身につけることを目標として，ロボットによる汎用的なスキル獲得を行うための研究開発を進めている．

Amazon：倉庫でのピッキングを完全自動化することを目標として，技術開発コンテストAmazon Picking Challengeを開催．

トヨタ自動車：自動走行実現のための基礎研究の成果として，深層強化学習により自ら動作を学ぶロボットカーを展示．

### 4.1.3 言葉の意味理解

言葉の意味理解に関わる取組は，基礎的な研究開発が始まり，着実に成果をあげているが，実用化には至っていない．例えば，Googleの研究チームは，ディープラーニングにより文意を理解して文章を要約することを目的としたテキスト要約の研究開発を進めている．

## 法律

### 4.2.1 知的財産

AIに関する知財の議論として，

1. **AIが創作に関わる「AI生成物」**
2. **学習済みモデル**
3. **学習用データ**

の知財制度上の扱いや活用と保護に向けた議論が行なわれている．以下では，上記３つに対する国内と海外の動向を示す．

#### 4.2.1.1 国内の動向

##### 4.2.1.1-① AI生成物

知的財産戦略本部により，AI生成物は，図２で示すように，AI利用者に創作的寄与の有無によって「**AIを道具として利用した創作**」と「**AI創作物**」に分類され，知財制度上，以下のように整理されている．

「AIを道具として利用した創作」：著作物性が認められ，著作者はAI利用者．

「AI創作物」：現行の著作権法上は著作物と認められない．

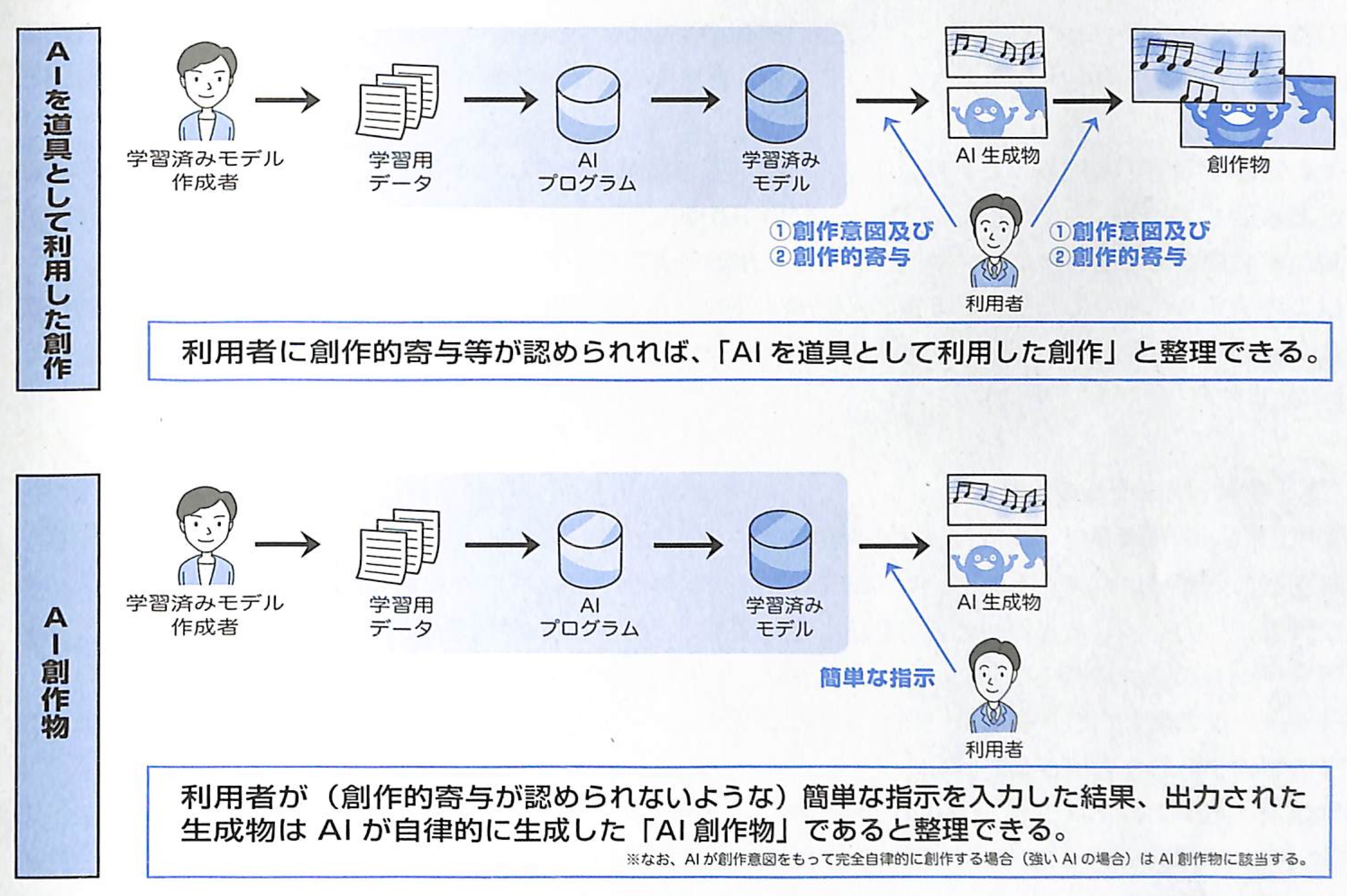


図 ：「AI生成物」の分類イメージ

##### 4.2.1.1-② 学習済みモデル

学習済みモデルは，次の２つの場合が考えられる．

・少なくともプログラムおよびパラメータと評価できる場合には，一般のプログラムと同様に特許法上の保護される．

・ネットワークの構造とパラメータが外から見えない場合でも，既存の学習済みモデルにデータの入出力を繰り返すことで得られる結果を基に学習することにより，同様のタスクを処理する別の学習済みモデルが効率的に作成できるもの（**蒸留モデル**と呼ばれる）は既存のモデルへの依拠性の立証が難しいので著作権による保護は困難である．したがって，特許権や契約による保護等の在り方について議論されている．

##### 4.2.1.1-③ 学習用データ

インターネット上のデータ等の著作物を元に学習用データを作成・解析することは，営利目的も含めて，著作権法47条の7に基づいて著作権侵害にあたらないとされている．この条項は，我が国独特のものである．

#### 4.2.1.2 海外の動向

##### 4.2.1.2-① AI生成物

諸外国においても，人間によって創作されたといえない以上，AI生成物は著作権保護を受けることはない．ただし，「コンピュータ生成物」について著作権を認めている英国は例外．

##### 4.2.1.2-② 学習済みモデル

諸外国においても，学習済みモデルが少なくともプログラムおよびパラメータと評価できる場合には，一般のプログラムと同様に特許法上の保護を受け得るとされている．

##### 4.2.1.2-③ 学習用データ

諸外国においても，我が国の著作権法47条の7に類似した法制度が，ドイツ，EUで整備されつつあるが，我が国と違って非営利目的に限られている．

#### 4.2.1.3今後の展望

今後の動向が注目されるAIと知財制度に対する議論

・現在の法律では著作権保護を受けられない，人間が創作したと評価できないAI生成物に対して， 何らかの法的保護を与えるべきではないかとの議論．

・我が国の著作権法47条の7は，機械学習の促進にとって有用だが，もともと機械学習を想定した規定ではないため，より機械学習の促進に資するように見直すべきではないかとの議論．

### 規制緩和・新たなルール

#### 4.2.2.1 自動運転

**自動運転レベルの定義**



**自動運転は，法規上許されるのか**

自動車に関する法整備でポイントとなるのは「**ジュネーブ道路交通条約**」と「道路交通法」である。ジュネーブ道路交通条約では，ドライバーの制御下にあることが前提とされている．したがって，現行の法規上「完全自動運転」は認められていない．自動運転について日本で現行認められるのは，運転者がいつでも操作できる状態にある「レベル３」までである．

**損害賠償責任について**

「レベル2」まで：刑事上の責任は原則運転手．

「レベル3」から：責任の所在はまだ明確になっていない．自動車メーカーやサプライヤーにまで責任が及ぶ可能性がある．

#### 4.2.2.2 ドローン

2015年12月施行の**改正航空法**（図３）により，ドローンの基本的な飛行ルールが定められた．簡単に述べると，次のようなルールである．

・飛行機と人が多い場所が、無許可では飛行禁止．

・ 逆に飛行機を邪魔せず、人口の少ない場所では、日中は許可無く飛行が可能．

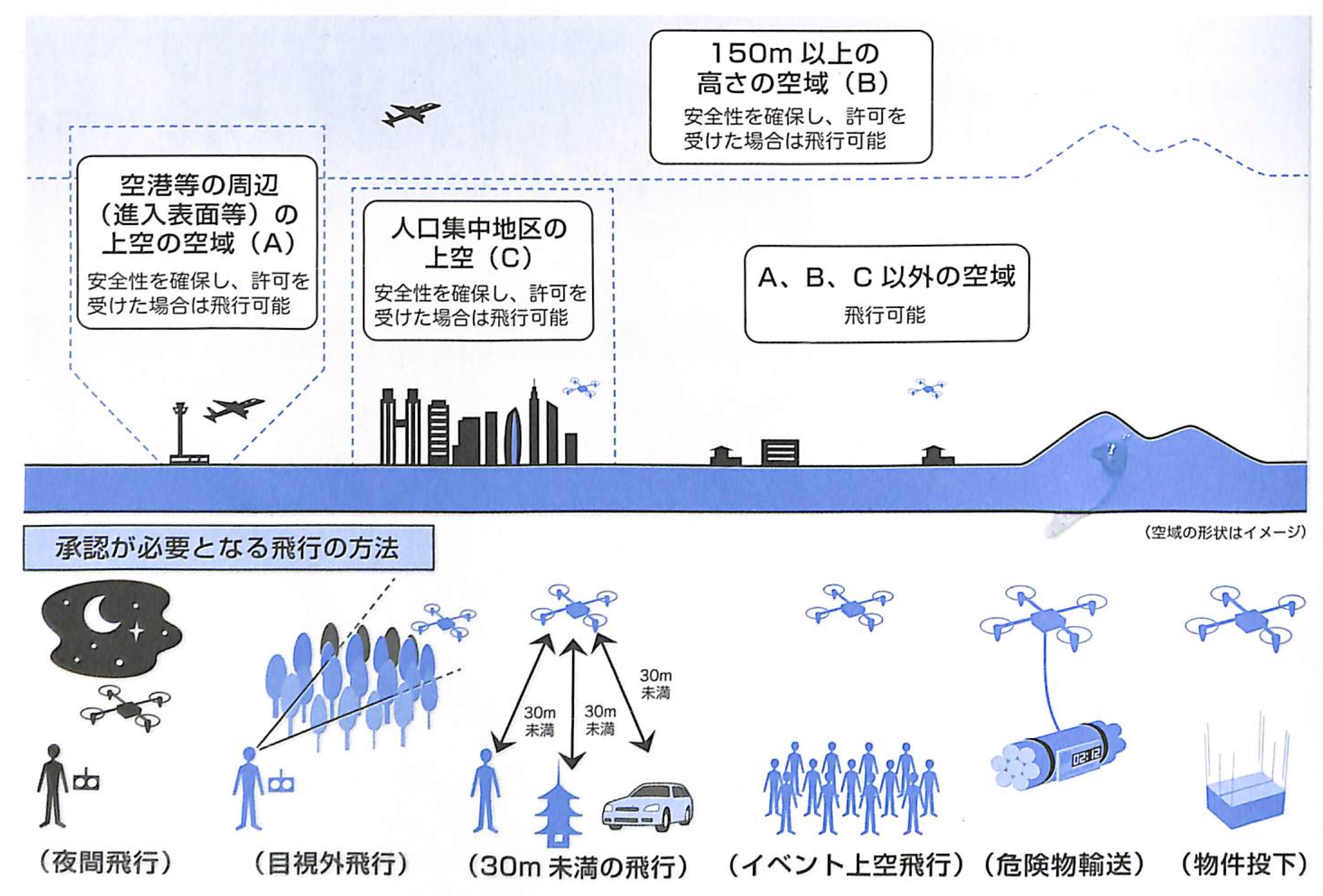


図 ：改正航空法における小型無人機の空域の考え方

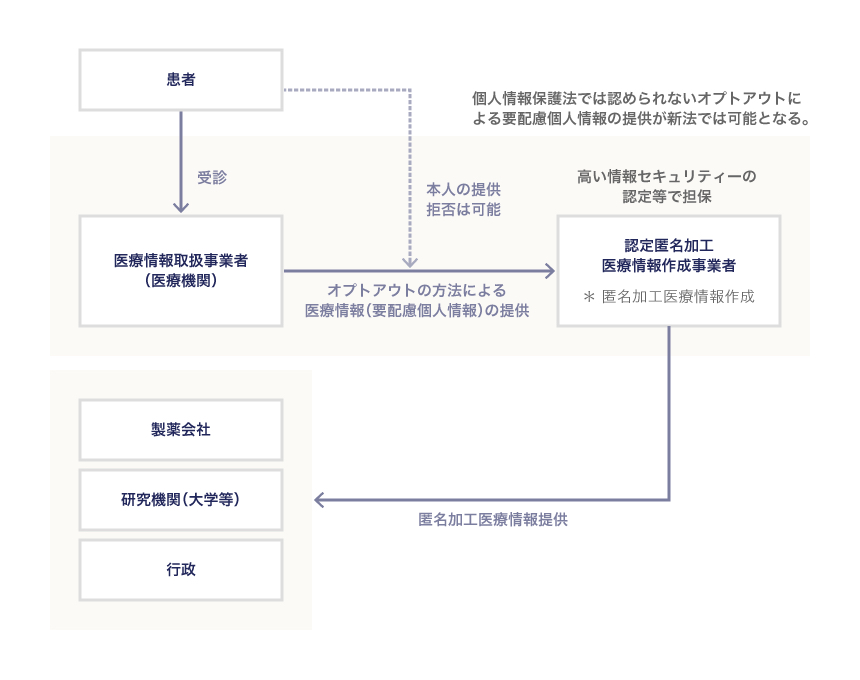
都市部の有人地帯においては，2020年代頃の本格運用に向けた制度の検討・整備がうたわれている．

#### 4.2.2.3 健康・医療・介護

ヘルスケア等の領域でのAI活用には，個人データを必要に応じて機関をまたいで連携・活用することが必要．これに対する取組は，近年始まったばかりだが，中でも重要なのは，**医療ビッグデータ法**こと、「**医療分野の研究開発に資するための匿名加工医療情報に関する法律**」．医療ビッグデータ法のポイントは，以下の２点である．

・医療情報（要配慮個人情報）のオプトアウトの手続きを認めること．

・認定匿名加工医療情報作成事業者の認定

オプトアウト：あらかじめ本人に対して個人データを第三者提供することについて通知・認識し得る状態にしておき，本人が反対をしない限り，同意したものとみなし，第三者提供をすることを認めること．

また，医療情報等の個人データの安全なデータ流通の仕組みとして，パーソナルデータストア（PDS）などの具体的な検討が進められている．

#### 物・サービスへのニーズとのマッチングや効率化

製造，流通，サービス業においては，AIを活用した事業者や業種を超えたカスタマイゼーションの実現が期待されている．そのため，個人情報保護に配慮したデータの利活用の新たな制度が検討されている．データの匿名化に関しては，2017年5月施行の改正個人情報保護法により，**匿名加工情報制度**が創設された．

## 4.3 倫理

### 4.3.1 背景にあるAIに関するリスク

|  |  |  |
| --- | --- | --- |
| リスク分類 | 具体的例 | 対応例 |
| 1. AI自身のリスク | AIの暴走 | AIへの理解の促進 |
| 1. AIに関わる人間のリスク | 悪意を持った人間がAIを利用 | AI関係者の倫理教育 |
| 1. 社会的負のインパクト | AIが雇用を奪う | AIでは対応できないスキルの整理，教育，訓練の充実 |
| 1. 法律・社会の在り方 | AIが起こした事故の責任の所在が曖昧 | AI実装時の法的問題に関する整理・議論・合意形成 |

AIには，大きく分けて以下の４つリスクがあると言われており，政府，民間，両方で問題の整理や対応が議論されている．

### 4.3.2 国内の取組

#### 6.2.1 政府

・**総合科学技術・イノベーション会議「人工知能と人間社会に対する懇談会」**

幅広い観点から，国内外の動向を俯瞰した上で，AIと人間社会の関わりについて取り組むべき課題や方向性を検討している．

・**総務省「AIネットワーク化検討会議」**

AIネットワーク化が社会・経済にもたらす影響とリスクの評価を行った上で，AIの開発原則・指針の策定などを整理している．

#### 民間

**・人工知能学会倫理委員会「人工知能学会　倫理指針」**

今後の人工知能学会と社会との対話に向けた大まかな指針を掲げる．

### 4.3.3 海外の取組

#### 4.3.3.1 政府

米国,　英国，EU政府では，以下のような問題が検討・考察・言及されている．

・AIの研究開発の透明性

・AIの普及が雇用に及ぼす影響と，それに対応した労働人口の育成の必要性

・ロボティクスとAIのプログラミング及び利用に起因する倫理的・法的問題

・AIをロボット技術が進展した場合の法律の在り方，人々の価値観，社会行動の変化

#### 4.3.3.2 民間

・**米国人工知能学会（AAAI）「Presidential Panel on Long-Term AI Futures:2008-2009」**

23名の研究者がAIの社会的な課題等について議論を行った．

・AIの予期せぬ行動に関連した責任割り当ての問題

・ロボットへの感情的な問題

・**スタンフォード大学「AI100」**

AAAIの議論を受けて2014年に立ち上げられた．以下の観点から法的・政策的な問題を整理している．

・プライバシー

・AIが関わる民事責任，刑事責任

・AIの人間に関する代理性

・AIの能力の認証

#### 4.3.3.3 国際的な枠組みでの議論

**IEEE「ETHICALLY ALIGNED DESIGN: A Vision for Prioritizing Human Wellbeing with Artificial Intelligence and Autonomous Systems」**

AIを人間の倫理的価値に沿ったものにするための指針のドラフト案を公開した．

# 文献目録

1. **岡谷貴之.** 深層学習 機械学習プロフェッショナルシリーズ. 出版地不明 : 講談社, 2015.

2. **ＡＩ白書編集委員会情報処理推進機構, (編).** ＡＩ白書　２０１７　人工知能がもたらす技術の革新と社会の変貌. 出版地不明 : 角川アスキー総合研究所, 2017.