

平成 2 9 年度
次世代医療機器・再生医療等製品
評価指標作成事業

人工知能分野
審査WG報告書

平成 3 0 年 3 月

審査WG座長 橋爪 誠

九州大学大学院 医学研究院
先端医療医学部門 先端医療医学講座

目 次

I. はしがき

II. WG 委員等名簿

III. 人工知能技術を利用した医用画像診断支援システムに関する評価指標案

IV. 委員報告

- 1：人工知能技術の医療応用：医療情報の観点から（横井 英人 委員）
- 2：新 AI 時代におけるコンピュータ支援診断（CAD）の新展開（藤田 広志 委員）
- 3：深層学習などの機械学習における課題（清水 昭伸 委員）
- 4：病理診断における人工知能による病理診断支援システムについて（佐々木 毅 委員）
- 5：消化器内視鏡領域における AI 技術応用（田中 聖人 委員）
- 6：超音波検査における人工知能の活用について（中田 典生 委員）
- 7：放射線診断分野における人工知能（待鳥 詔洋 委員）

V. 参考資料

- 1：平成 23 年 12 月 7 日付け薬食機発 1207 第 1 号厚生労働省医薬食品局審査管理課
医療機器審査管理室長通知「次世代医療機器評価指標の公表について」別添 3
「コンピュータ診断支援装置に関する評価指標」
- 2：第 1 回会議 PMDA 講演資料「医療機器の承認審査について」
- 3：第 1 回会議 厚生労働省講演資料「人工知能を利用した医療機器の承認審査に
おける課題 ～評価指標の位置づけと期待～」
- 4：第 1 回会議 東京大学 松尾 豊先生講演資料「人工知能は人間を超えるか
ーディープラーニングの先にあるもの」
- 5：第 3 回会議 個人情報保護法委員会提供資料「個人情報保護法の概要」
- 6：第 3 回会議 内閣官房 健康・医療戦略室提供資料「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について」
- 7：第 18 回次世代医療機器・再生医療等製品評価指標検討会／医療機器開発ガイドライ
ン評価検討会審査 WG 活動報告資料

VI. 平成 29 年度 WG 会議議事概要

I. はしがき

<はしがき>

九州大学大学院医学研究院 先端医療医学部門

橋爪 誠

コンピュータ診断支援（Computer-Aided Diagnosis:以下 CAD）とは、X線画像に代表される放射線画像をはじめとする医用画像に対して、コンピュータで定量的に解析された結果を「第2の意見」として利用する「医師による診断の支援」である。

当該診断支援を行う医療機器に関しては、「次世代医療機器評価指標の公表について」（平成23年12月7日付け薬食機発1207第1号厚生労働省医薬食品局審査管理課医療機器審査管理室長通知）別添3「コンピュータ診断支援装置に関する評価指標」にて公表しているところであるが、近年の計算機技術の著しい発展により、学習データのさらなる追加（事後学習）等により市販後に性能が変化する人工知能技術を利用した診断支援装置については述べられてない。また、人工知能技術の中でも、近年着目されている深層学習（Deep Learning）では、最終結果を導き出すアルゴリズムがブラックボックス化している点が特徴であると同時に、事後学習等による性能変化の内容も基本的にはアウトプットでしか評価できないこと等から、その評価や運用に当たって従来にない新たな課題が生じることが指摘されている。当該診断支援装置を用いた際、学習データに含まれていなかったタイプの事例に対する性能をどのように評価するのか（又はそのような事例にどのように対応するのか）、学習データの真正性や偏り等を考慮してどのように学習の品質確保を行うのか、さらに、市販後の性能変化に対する評価方法等に関して、従来とは異なる新しい視点も考慮する必要がある。

このような状況に鑑み、本評価指標は、上記通知も踏まえて、人工知能技術を利用した医用画像診断支援システムの有効性と安全性を評価する際の問題点や留意すべき点について、現時点での考え方をまとめたものである。

その結果、1）データの信頼性に関しては、審査時にデータの取得方法について申告すべきである。2）市販後に性能が変化することにより、アウトプットが患者に良い結果を及ぼす場合と害を及ぼす場合とが考えられる。現時点では、性能が承認時と大きく変化する医療機器の市販化は、技術的に対応がまだ十分でなく、企業および審査側とも慎重な姿勢であることがわかった。しかし、海外ではこれらの医療機器をAIのもつ特性として受け入れて、市販化の傾向にあり、海外調査をするなど継続して検討する必要があると考えられた。

今後の対応策として、(1) 性能変化を自ら管理できる機能や、機能をリセットさせ承認時の機能に戻すようにしておく。(2) 審査時に性能が変化する許容範囲について予め承認を得ておく。(3) 許容範囲を逸脱する場合には、別製品とみなし、再度審査を受ける必要がある。(4) 性能の変化については、記録を残し、事後の検証や、今後の開発推進に役立てる必要がある。(5) 基本的に、診断の最終責任は医師である。などが議論された。

Ⅱ．人工知能分野審査 WG 委員等名簿

人工知能分野審査WG委員等名簿（敬称略）

座長

橋爪 誠 九州大学大学院 医学研究院 先端医療医学部門 先端医療医学講座 教授

委員（五十音順）

佐々木 毅 東京大学 医学部 人体病理学・病理診断学分野 准教授（日本病理学会推薦）

清水 昭伸 東京農工大学大学院 工学研究院 教授

田中 聖人 京都第二赤十字病院 医療情報室 室長（日本消化器内視鏡学会推薦）

中田 典生 東京慈恵会医科大学 放射線医学講座 准教授（日本医用画像工学会推薦）

藤田 広志 岐阜大学 工学部 電気電子・情報工学科 情報コース 教授
（医用画像情報学会推薦）

待鳥 詔洋 国立研究開発法人 国立国際医療研究センター 国府台病院 放射線科診療科長
（日本医学放射線学会推薦）

横井 英人 香川大学医学部附属病院 医療情報部 教授

厚生労働省

中井 清人 医薬・生活衛生局 医療機器審査管理課 課長

柳沼 宏 医薬・生活衛生局 医療機器審査管理課 課長補佐

青柳 ゆみ子 医薬・生活衛生局 医療機器審査管理課 医療機器規制国際調整官

独立行政法人 医薬品医療機器総合機構（PMDA）

高江 慎一 医療機器審査第一部 部長

岡崎 譲 医療機器審査第一部 審査役

菅原 明彦 医療機器審査第一部 審査役補佐

加藤 健太郎 医療機器審査第一部 審査専門員

佐々木 拓哉 医療機器審査第二部 審査役補佐

桐山 瑤子 医療機器審査第二部 審査専門員

古森 亜矢 医療機器審査第二部 嘱託

谷城 博幸 医療機器審査第三部 審査役

宮崎 生子 規格基準部 部長

今川 邦樹 規格基準部 医療機器基準課 基準専門員

長谷川 浩一 科学委員会事務局

若尾 りか 科学委員会事務局

国立医薬品食品衛生研究所（事務局）

齋島 由二 医療機器部 部長
中岡 竜介 医療機器部 室長
加藤 玲子 医療機器部 主任研究官

産業技術総合研究所（オブザーバー）

鎮西 清行 ヒューマンライフテクノロジー研究部門 副研究部門長
坂無 英徳 人工知能研究センター 人工知能応用研究チーム 主任研究員

Ⅲ．人工知能分野を利用した医用画像診断支援 システムに関する評価指標案

人工知能技術を利用した医用画像診断支援システムに関する評価指標（案）

1. はじめに

コンピュータ診断支援（Computer-Aided Diagnosis:以下 CAD）とは、X線画像に代表される放射線画像をはじめとする医用画像に対して、コンピュータで定量的に解析された結果を「第2の意見」として利用する「医師による診断の支援」である。

当該診断支援を行う医療機器に関しては、「次世代医療機器評価指標の公表について」（平成23年12月7日付け薬食機発1207第1号厚生労働省医薬食品局審査管理課医療機器審査管理室長通知）別添3「コンピュータ診断支援装置に関する評価指標」にて公表しているところであるが、近年の計算機技術の著しい発展により、学習データのさらなる追加（事後学習）等により市販後に性能が変化する人工知能技術を利用した診断支援装置については述べられていない。また、人工知能技術の中でも、近年着目されている深層学習では、最終結果を導き出すアルゴリズムがブラックボックス化している点が特徴であると同時に、事後学習等による性能変化の内容も基本的にはアウトプットでしか評価できないこと等から、その評価や運用に当たって従来にない新たな課題が生じることが指摘されている。当該診断支援装置を用いた際、学習データに含まれていなかったタイプの事例に対する性能をどのように評価するのか（又はそのような事例にどのように対応するのか）、学習データの真正性や偏り等を考慮してどのように学習の品質確保を行うのか、さらに、市販後の性能変化に対する評価方法等に関して、従来とは異なる新しい視点も考慮する必要がある。

このような状況に鑑み、本評価指標は、上記通知も踏まえて、人工知能技術を利用した医用画像診断支援システムの有効性と安全性を評価する際の問題点や留意すべき点について、現時点での考え方をまとめたものである。

2. 用語の定義・説明

本評価指標における用語の定義等は以下のとおりである。

（1）医用画像診断支援システム

臨床現場において、医師が様々な画像撮影装置を用いて画像診断（読影）を行う際に参考となる情報を提供するシステム又はソフトウェアであり、CAdEやCAdxの機能を持つ。支援の方式としては使用者（医用画像診断支援システムを実際に臨床現場で使用する医師等）に対する支援形式に応じ、first reader、second reader、con-current reader等が存在する。

1）CAdE (Computer-Aided Detection)

画像上で病変の疑いのある部位をコンピュータが自動検出し、その位置をマーキング

する機能を有する単体ソフトウェア又は当該ソフトウェアが組み込まれている装置。コンピュータにより医用画像データのみ又は医用画像データと検査データの両方を処理し、病変又は異常値の検出を支援する。

2) CADx (Computer-Aided Diagnosis)

病変の疑いのある部位の検出に加え、病変候補に関する良悪性鑑別や疾病の進行度等の定量的なデータを数値やグラフ等として出力する機能を有する単体ソフトウェア又は当該ソフトウェアが組み込まれている装置。診断結果の候補やリスク評価に関する情報等を提供等により診断支援を行うものを含む。

注) CADe と CADx の定義は前出の通知より一部修正の上、引用

(2) 人工知能

人の高度な知能によって行われている推論、学習等を模倣するコンピュータシステム又はソフトウェア。本評価指標においては、特に、臨床で使用を開始した後、追加データによる機械学習（例えば、深層学習等）により、その性能が変化する機能（事後学習機能）を有するものを対象とする。

(3) 機械学習

人工知能技術の一分野として大きく発展した技術であり、人間が行っている学習等の機能をコンピュータ（ソフトウェア）で実現するための手法。アルゴリズムはさまざまであり、判別分析法等の古典的手法から、決定木、ニューラルネットワーク、サポートベクターマシン等、多数の方法が提案されている。代表的な応用としてはクラス分類や回帰があるが、これらは CAD の目的と一致していることから、多くの CAD システム開発時に利用されてきている。

(4) 深層学習

機械学習の一つであり、最近特に注目されている学習方法。多層化したニューラルネットワークを大量のデータを用いて学習させることに特徴がある。ニューラルネットワークには、階層型ネットワークのような確定的モデルと、ボルツマンマシンのような確率的モデルがある。学習アルゴリズムには、教師あり、教師なし、半教師学習以外に強化学習等もある。

2006 年の Hinton らの研究に始まり、囲碁の世界王者を倒した AlphaGo や、各種コンペティションにおいて上位を占めたソフトウェアで使用され、当時の常識を塗り替えるほど突出した成績を示したことから盛んに研究されている。性能の飛躍的向上の一つの理由は、多層のネットワークを学習可能とする新しい技術の登場である。また、深層ニューラルネットワークの学習のための大量のデータと並列分散計算環境の整備もある。

（５）転移学習

20 年以上も前から認識されている機械学習の一分野であり、「新規タスクの効果的な仮説を効率的に見つけ出すために、一つ以上の別のタスクで学習された知識を得て、それを適用する問題」とされている。

近年、深層学習でも盛んに転移学習が用いられるようになったが、その具体的な利用例として以下が挙げられる。

１）大規模なデータセットで学習させたニューラルネットワークを特徴抽出器として利用し、得られた特徴を別のタスクへ転用する。例えば、一般の自然画像で学習済みのニューラルネットワークに医用画像を入力し、中間層の出力を特徴量としてその医用画像の認識・診断支援などに利用する場合が考えられる。

２）別の目的で学習済みのニューラルネットワークを用意し、異なるタスクの学習データを用いて再学習させる。例えば、CT 像の認識・診断支援を目的とする場合、１）に例示した自然画像で学習済みのニューラルネットワークに CT 像と教師データを与えて再学習させる場合が考えられる。この場合には、一からニューラルネットワークを学習させる場合と比べて、少数の CT 像と教師データで性能の高いニューラルネットワークが得られることが報告されている。

3. 本評価指標の対象

本評価指標は、医用画像診断支援システムのうち、人工知能技術を利用して様々な画像撮影装置で得られた臨床画像から、

- ・ 疾患名まで特定せず、病変の疑いがある部位のみ検出する（いわゆる CAdE）
- ・ 病変の疑いがある部位の検出に加え、疾患名の候補を提示する（いわゆる CAdx）
- ・ 疾患名の候補の提示に加え、それらの重み付け（順位づけ）まで行う（同上）

等の機能を有するシステム全て（CAdE、CAdx 等、以降「支援システム等」と記述する。）を対象とする。

対象とする支援システム等は、人工知能技術の利用の形態や学習アルゴリズムの種類にかかわらず、機械学習等により市販後も事後学習等に伴い診断支援性能が変化することを意図したものとする。なお、対象はあくまで診断を行う医師を支援するシステムであり、仮にシステムが想定外の挙動、誤動作をした場合に、使用者側でそれを検知できることが基本要件となる。

事後学習により生じる性能変化に伴う品質管理を製造販売業者が行うことが難しい人工知能技術を利用した支援システム等、例えば、使用者が事後学習させることにより施設ごとに異なる性能変化が生じ得るものに関しては、解決すべき事項が数多く存在することか

ら、本評価指標における対象とすることは困難である。しかしながら、人工知能技術の著しい発展に伴い、これらの支援システム等が近い将来開発される可能性に鑑み、それらの品質、安全性及び有効性を確保するための基本的な考え方を別添に取りまとめた。

なお、開発する支援システム等が本評価指標に該当するか判断し難い場合は、必要に応じ厚生労働省医薬・生活衛生局医療機器審査管理課に相談すること。

4. 本評価指標の位置付け

近年、支援システム等の開発では、深層学習のようにデータから自動で学習をする機械学習の利用が主流になっている。特に、大量のデータを継続的に収集する仕組みが整備され、学習に用いるコンピュータの性能が飛躍的に向上したことにより、短期間で繰り返し学習をすることが可能となった。また、臨床で得られたデータや支援システム等の結果を自動的に収集し、事後学習データとして人工知能に与えて性能を変化させることも技術的には実現可能となった。そのため、臨床で使用しながらその性能が段階的又は連続的に変化する支援システム等の実現が期待されている^{注)}。

本評価指標は、対象とする支援システム等において、現時点で考えられ得る問題点、留意すべき事項を示したものであるが、今後の更なる技術革新や知見の集積等を踏まえ改訂が必要なものであり、承認申請内容等に関して拘束力を有するものではない。

支援システム等の評価に関しては、個別の人工知能構築方法、使用目的及び製品特性を十分に理解した上で、科学的な合理性をもって、柔軟に対応することが必要である。なお、本評価指標以外に前出の「コンピュータ診断支援装置に関する評価指標」及び現存する国内外の関連ガイドライン等、例えば、平成 28 年 3 月 31 日付け厚生労働省医薬・生活衛生局医療機器・再生医療等製品担当参事官室事務連絡「医療機器プログラムの承認申請に関するガイダンスの公表について」別添「医療機器プログラムの承認申請に関するガイダンス」等を参考にすることも考慮すべきである。

注) 深層学習の事後学習の対象には、ネットワークの結線（結合）の重み以外に、ネットワークのハイパーパラメータ（層数、特徴マップ数、畳み込みのフィルタサイズ、ユニット数等）も考えられる。これらのハイパーパラメータも含めた学習とは、例えば、パラメータの組み合わせごとに教師データとネットワークの出力間の誤差（損失）を最小化する学習を網羅的に行い、それらの中から誤差最小のハイパーパラメータを選択すること等が挙げられる。ただし、変更されたハイパーパラメータの程度によっては、軽微な変更ではなく支援システム自身の変更と判断される可能性があるため、当該変更の実施については、事前に独立行政法人医薬品医療機器総合機構と相談することを推奨する。

5. 支援システム等で想定される主な課題と解決の方向性

(1) ブラックボックスとしての性質

機械学習、特に深層学習は、その特性上、ニューラルネットワークによる判断の過程がブラックボックスであり、アウトプットの判断根拠を明確に示すことができない。このため、通常の医療機器ではその性能確保のために示すことが求められる動作原理の明示が困難となる。そのため、診断アルゴリズムを詳細に精査するというよりは、一般的な医療機器プログラムの評価と同様、そのインプットに対して所要のアウトプットが得られているかを確認することに重点を置き、その性能を評価することが適当と考えられる。

また、動作原理及びアルゴリズムの詳細を記述することが困難であるため、性能の確認が行われた事例以外（例えば、開発時の学習データセットに含まれていなかった希少事例）での挙動等が想定しにくいことも考えられる。このようなリスクに鑑み、機械学習、特に深層学習を利用した支援システム等においては、その性能を確保するために、使用方法及び目的に応じて、製造販売業者が対象疾患の検出率、偽陽性率、偽陰性率、検出に要する時間等、性能に係る値を適切に規定し、その性能が常に担保されていることを示すことが必要となる（6（3）項を参照）。

支援システム等が想定外の挙動を示した場合に、使用者に対して当該挙動を通知する仕組みも必要となる。例えば、想定外の挙動が判断可能な使用者のみに使用を許可する等で使用者側に一定の要件を求める、又は臨床での使用状況を逐次フィードバックすることで、製造販売業者側が性能検証可能となる仕組みを構築する等、支援システム等の目的に応じた適切な方策を用意することが望ましい。（6（4）項を参照）。

想定外の挙動、誤判定等を完全に無くすことは極めて難しいが、そのような情報を収集し、対策を講じることができるような運用方法もあらかじめ考慮しておく必要がある。

(2) 性能変化

支援システム等においては、人工知能技術を利用することで事後学習による診断アルゴリズムの変化を受けてその性能が向上することが期待される反面、学習の内容等によっては意図しない性能変化が生じ、当初規定していた性能を下回る可能性も否定できない。このような利点とリスクとを勘案し、支援システム等には、継続的な性能検証によるほか、学習用データの要件や学習のプロセス等を規定し、学習の品質を確保するための対応策が求められることになる。

1) 継続的な性能の検証方法

支援システム等の品質、安全性及び有効性を確保する上で、性能が変化するたびに検

証する必要がある。臨床使用に伴って事後学習が行われ、性能が連続的又は相当程度高い頻度で変化する場合には、特にその内容、プロセス管理に加え、変化した性能の検証及びその管理が重要となる。このことを踏まえ、あらかじめ市販後に生じ得る性能変化水準を臨床上許容される範囲かつ統計学的な妥当性に基づいた範囲で目的に応じて規定しておくこと等とともに、その機器に適用可能な検証方法、変化した性能が規定された範囲を逸脱した際の対策等を定めておくことが求められる（6（3）、6（4）項を参照）。また、検証に使用したデータについては、入手元を明らかにした上でその使用根拠や妥当性を示すことが求められる（6（2）2）項を参照）。

2）性能変化に伴う品質確保

人工知能の性能向上には、その人工知能の学習方式に適した事後学習が必須であることから、その学習方式を明確にした上で、使用した事後学習データの詳細を示すことが求められる。（6（2）2）項を参照）。

支援システム等においては、まず、一定量蓄積した臨床データを事後学習データとして使用することで性能が段階的に変化することが考えられることから、基本的にその性能変化のための学習及び性能検証は製造販売業者が実施することとなり、市販前に求められる要件と同様に性能変化後の品質を確保することが求められる（6（3）1）項を参照）。

一方、事後学習により性能が連続的又は相当程度高い頻度で変化する機器の場合、製造販売業者の意図に反してその性能が低下し、臨床上許容される水準を下回る可能性がある。このような問題を防ぐために、原則、人工知能は製造販売業者が管理した上でネットワークを介して支援システム等と接続する形式とするとともに、当該支援システム等の目的及びリスクを勘案した検証方法や対策を講じることが求められる（6（3）2）②項を参照）。

性能が変化した支援システム等の品質確保に係る作業は製造販売業者が実施することとするが、品質確保の目的で、使用者による事後学習機能を有さない、誤使用や適応外使用を防止するため使用者を対象としたトレーニングを実施する、性能が下限以下となった場合の対策を使用者側の端末にも施す等、その支援システム等のリスクに応じて必要な追加対策を施すことが望ましい（6（4）項を参照）。

3）薬事上の手続きにおける考え方

一般的には支援システム等の性能を向上させた結果、臨床上の有効性が向上する場合、薬事上の手続きが必要になる場合があるが、その要求範囲は性能変化の内容及びその変化に伴うリスクを勘案した設定が必要となる。

性能向上に係るバージョンアップが新規製品となる、又はその適用に費用が発生する

可能性があることから、その性能変化は対象となり得る販売後の製品全てに適用されず、同じ製品であってもバージョンの異なるものが市場に存在することになる可能性が高い。このような、人工知能の流通体系の特殊性を考慮しながら、バージョンアップに関するフォローアップ体制や薬事上の効率的な手続きについても検討が必要になる。

（３）責任の所在

人工知能を利用した医療機器においては、その使用に伴う責任の所在が曖昧になることが想定される。特に、支援システム等においては、その目的が「診断支援」であるにもかかわらず、医師が最終診断まで委ねてしまう危険性を含んでいる。よって、対象となる機器の品質、安全性及び有効性の確保、適正使用のために行うべき方策（使用者に対するトレーニング等）、保守、トラブル対応においては製造販売業者が責任を負うこと、その機器の適切な使用及び医療行為に伴う有害事象（支援システム等における設計・仕様上の不具合に起因した有害事象を除く。）等に関しては医師が責任を負うことを明確にする必要がある（６（４）項を参照）。

6. 評価に際して留意すべき事項

（１）基本的考え方

本評価指標で対象とする支援システム等は、その特性上、通常の医療機器と異なり、原理（実装する検出・診断アルゴリズム等）や設計仕様等のみで性能を確保することが難しいことから、性能に影響する要素、効果を確認できる範囲や限界等を規定する必要がある。当該要求事項を達成するためには、医学的かつ統計学的に妥当な方法で、それらを規定するとともに支援システム等の性能を検証することが求められる。なお、支援システム等自体の品質、安全性及び有効性を確保するために必要となる評価は、原則として前出の「コンピュータ診断支援装置に関する評価指標」を参考に、目的に応じた性能水準を設定の上、実施することになる。

上述した内容を考慮した上で、現時点の科学水準を基に、本評価指標の対象となる支援システム等の評価において特に留意すべき点を示す。

（２）品目の検出・診断の原理、学習、情報セキュリティ等に関する基本的事項

１）検出・診断の原理（アルゴリズム）

具体的にその記述が可能な検出・診断アルゴリズムにおいては、従前通り、そのアルゴリズムとプログラム概要を示すことが求められる。深層学習のように処理工程がブラックボックス化する原理を用い、承認申請時に検出・診断アルゴリズムを示すことが困難な場合には、設計開発時の検出・診断用ネットワーク構造とプログラム概要を示す必要がある。なお、本評価指標が公表された後、ブラックボックス化したネットワーク構

造等の振る舞いが説明できるようになった場合には、可能な限りその情報を示すことが求められる。

2) 学習

本評価指標が対象とする支援システム等は、適切な学習データを使用した人工知能の学習によりその目的を達成するために必要となる性能を有することが求められる。よって、対象とする支援システム等の機構や規定された性能等に応じて、以下の項目例を参考にして明確にすべき項目、それらを使用した根拠及び妥当性を示す必要がある。

- ・ 学習アルゴリズムとプログラム概要
(教師あり、半教師あり、強化学習、自己学習、その他等を明確にすること)
- ・ データ (学習データ、バリデーションデータ^{注1)} 及びテストデータ^{注2)} について、以下を参考に必要な項目を明記すること)

注1) バリデーションデータ：機械学習アルゴリズムのハイパーパラメータ (例：サポートベクターマシンの目的関数を決定するパラメータ、深層学習のネットワークの層数、特徴マップ数、畳み込みのフィルタサイズ、学習回数等) を決定するためのデータ

注2) テストデータ：システムの性能を評価・検証するためのデータ

- データソース (取得及び管理方法の明確化を含む。特に、支援システム等により取得した画像等を事後学習データとする場合には、対象となる患者から同意を取得する方法も含む。)
- 画像データ取得時の撮影パラメータ
- 紐付けした臨床データの種類^{注3)} (アノテーション (例：病変等の位置やサイズ (ラベル画像も含む。)、良悪性の鑑別結果) も含む。)
- 臨床画像データ以外のデータ使用の有無^{注4)} (Data Augmentation 等の目的で作成したコンピュータシミュレーションやファントム使用による画像データ)
- 画像データの枚数、サイズ、濃度レベル数等
- 学習前に加工 (例：匿名化) している場合にはその処理方法
- 臨床データの最終判断を行った医師等
- 臨床画像データ以外のデータを使用した場合は、その最終判断を行った責任者等
- テストデータを学習のプロセスから完全に切り離して管理するための方策

注3) これまでの医用画像診断支援システムは医用画像のみを用い、臨床的信頼度が高いアノテーションを用いて設計することがほとんどであった。しかし、技術の進歩により、近年では、学習データとして自然画像や人工画像、さら

には紐づけされた臨床情報の信頼性が低いデータ（Weak label データ）を用いることが可能となってきたが、このようなデータを利用する場合は、その使用根拠及び妥当性を示す必要がある。また、テストデータに関する各種留意点に関しては、その評価段階に応じて6（3）項を参照すること。

注4）特に転移学習の場合には、自然画像や別の医用画像、さらには人工画像等を学習データとして用いることが想定される。自然画像を用いた場合には画像中の対象部分、種類及び枚数等、別の医用画像を用いた場合には通常の医用画像を使用した場合と同等の事項、人工画像の場合にはその作成方法等について、その詳細を示すこと。

学習において使用する画像等をデータベース等から入手して使用する場合は、その妥当性を確認することが求められる。以下に、データベース等に関して承認申請時に記載が必要と考えられる点を例示する。なお、記載すべき項目は、データベースごとに異なるため、適切に選択、追加すること。

- ・ データベース運営者（学会、認定匿名加工医療情報作成業者等）の概要、組織体制等
- ・ 運営者が所有する事業計画書
- ・ 運営者が外部に委託している業務内容
- ・ データベースに保有するデータの種類
- ・ データベースの概要及び設計書
- ・ データ管理に係る各種手順書とその運用状況 等

上記の学習にかかわる全ての項目については、設計開発時の学習と事後学習で共通の項目、相違する項目を区別して明記する必要がある。例えば、設計開発時の学習と事後学習で学習アルゴリズムが異なる場合は、それぞれについて明らかにする必要がある。また、事後学習により機械学習アルゴリズムのハイパーパラメータも再決定される場合は、変化しうる範囲をあらかじめ規定したうえで、そのハイパーパラメータの変化が規定範囲内となることを明らかにすることが求められる。さらに、事後学習後のネットワークの変遷については性能変化ごとに必要な記録を保存し、後日検証可能な状態にすることが必要となる。

3）使用環境及び情報セキュリティ

人工知能技術の利用において、その技術がコンピュータ上にあるのか又はネットワークを介して接続しているのかを明確にした上で、次の要件を示すことが求められる。

- ・ 画像撮影装置からのデータ転送方法
- ・ 使用に必要なソフトウェアのインストール等による画像撮影装置又はその他のソ

ソフトウェアへの干渉

- ・ 支援システム等がインターネットと接続されている場合、平成 27 年 4 月 28 日付け薬食機参発 0428 第 1 号・薬食安発 0428 第 1 号厚生労働省大臣官房参事官（医療機器・再生医療等製品審査管理担当）、厚生労働省医薬食品局安全対策課長連名通知「医療機器におけるサイバーセキュリティの確保について」を参考にサイバーセキュリティ対策に必要な措置を講じていること
- ・ 支援システム等を使用して得られた臨床情報を事後学習データとする場合は、以下のガイダンスを参考に適切な管理を行い個人情報の保護を行うこと
 - 「医療・介護関係事業者における個人情報の適切な取り扱いのためのガイダンス」平成 29 年 4 月 14 日、個人情報保護委員会、厚生労働省
 - 「医療情報システムの安全管理に関するガイドライン第 5 版」平成 29 年 5 月、厚生労働省

注）必要に応じて、個人情報保護法及び次世代医療基盤法も参考にすること

4）併用する画像撮影装置等の条件

支援システム等が所定の機能を達成するにあたっては、あらかじめ併用する画像撮影装置で得られた医用画像データのフォーマットや解像度等が当該システム等の要求を満たす必要がある。また、市販されている特定の画像撮影装置全てに適用することが可能なものであっても、支援システム等の性能は、人工知能の学習に使用する医用画像データの撮像条件や教師データの作成方法等に影響されるため、必要に応じてそれらについて規定する必要がある。

（3）安全性、品質及び性能評価

1）市販前

支援システム等の評価においては、対象となる機器の目的やリスクに応じた評価が必要ため、品質、安全性及び有効性の評価に必要な規格値（検出率等）を明確にするとともに、その妥当性を示した上で適切な評価を行うことが必要である。

各種評価においては、6（2）項に記載した項目以外に支援システム等に特有の問題として、

- ・ 接続する画像撮影装置の機能（特に既存の画像撮影装置を利用可能なシステムの場合）
- ・ 責任者を含めた性能検証方法の明確化（治験による性能検証の場合は、「正解」となる最終的な確定診断を得るまでのプロトコル、後向き検証の場合は検証用テストデータに関して、その妥当性を示すために必要な情報を示すことが求められる）

- ・ 性能検証に使用したテストデータの種類、入手元及び妥当性（必要に応じて管理方法等も明らかにすること）
- ・ 市販後に生じ得る性能変化の範囲の規定（検出率の下限、偽陽性率及び偽陰性率の上限設定等）及びその規定根拠となる医学的又は統計学的データ等
- ・ 市販後に性能変化が発生しても、有効性及び安全性が確保されていることを確認するための方策（学習させるデータの入手方法、変化後の性能の検証方法等）
- ・ 性能変化に伴って生じ得る不具合

等、その目的に応じて考慮すべき項目を明確にする必要がある。

特に、性能検証用のテストデータセットは、学習データやバリデーションデータとは重複しないことが求められる一方、対象とする母集団の特徴を考慮して、十分な説明が可能となる質と量のデータが必要となる。しかしながら、機械学習の方法、目的及び対象に応じてその必要数は異なり一律に定義することは困難であるため、支援システム等の性能評価に使用したデータセット及び方法論を記載するとともにその妥当性検証スキーム等を個別に記載することが求められる。

2) 市販後

市販後、支援システム等に発生する性能変化に関しては、承認申請時に規定された性能検証方法に従い、事後学習データにより変化した性能等が臨床上又は統計学的な観点からあらかじめ規定された範囲内を満たし、品質が確保できることを確認の上、その妥当性検証を行うことが求められる。

支援システム等が既存の画像撮影装置に接続して使用可能な場合、画像撮影装置の性能に依存して入手できる画像データの質が多様となることが想定される。この場合には、得られた画像データが学習データとして使用できることを示す必要があり、その条件を満たさないデータは学習データとしないための対策が必要となる。また、性能検証及びそのために用いるテストデータセットに関する考え方は、基本的に前項と変わらない。加えて、事後学習データ及びテストデータの臨床最終判断を行った医師等を明確にする必要がある点も、前項と変わらない。

なお、人工知能にネットワークを介して接続する支援システム等において、使用者に必要な性能及び動作の詳細を漏れなく把握できるよう、性能変化が生じた際には使用者にその旨を通知するシステムを有することが望ましい。

以下に、性能変化の実現方法ごとに必要となる留意点を示す。

① バージョンアップごとの段階的な性能変化を意図する場合

製造販売業者が臨床現場で入手したデータ等を定期的に収集し、データの偏り等にも配慮しながら適切な形で人工知能に事後学習させることにより支援システム等の

性能を変化させる場合がこれに該当する。なお、臨床現場で入手したデータ以外を利用することも想定されることから、それらの利用においては承認申請時に明記するとともにその妥当性を示すことが求められる。

この場合、変化した性能の評価や、その変化に伴う安全性及び品質の確保は製造販売業者によってバージョンアップごとに行うこととなるが、それらの評価にあたっては承認申請時と同様に前出の「コンピュータ診断支援装置に関する評価指標」を参考に行うこと。

また、支援システム等の性能変化において必要となる薬事上の手続きについては、システムの特性にも依存するため、想定可能な範囲については独立行政法人医薬品医療機器総合機構と承認申請前に相談して決定しておくことが望ましい。

② 臨床使用に伴った事後学習により連続的又は相当程度高い頻度でのバージョンアップを意図する場合

支援システム等を臨床現場で利用して入手したデータを事後学習データとして使用し、連続的又は相当程度高い頻度で性能を変化させる場合には、使用する人工知能は製造販売業者が一括して管理し、各支援システム等とネットワークを介して接続していることが想定される。この場合、自動的に追加されたデータを用いた事後学習により変化した支援システム等の性能や品質があらかじめ規定された範囲に収まることから、その学習プロセス及びその管理方法を示すことが求められる。加えて、性能が規定範囲に収まることを検証するための機構をあらかじめ付与するとともに、製造販売業者で定期的にその性能検証を実施することが求められる。なお、当該バージョンアップの場合には、想定外の挙動、誤判定等の問題が生じる可能性が高いため、その解決策(リスクマネジメント)を講じておくことが求められる。

(4) リスクマネジメント

1) 基本的考え方

支援システム等におけるリスクとしては、事後学習の結果性能が変化し「人工知能が規定された変化範囲を逸脱した結果、誤った情報を提示すること」が挙げられる。バージョンアップごとの段階的な性能変化を意図する支援システム等の場合には、その性能変化によって生じた問題は製造販売業者が解決してから上市することが前提となるため、そのリスクは低い。一方、事後学習データが直ちに実装されることで性能が連続的又は相当程度高い頻度で変化する支援システム等の場合には、上記リスクが生じる可能性があるため、そのリスクを回避するためのシステムが必要となる。

また、適応外使用に伴い、上記リスクに加え、使用者が支援システム等における想定外の挙動に気付かない可能性も増大することが考えられるため、適応外使用を防ぐ方法

を講じることが求められる。

2) 対策について

対策が必要となる支援システム等には、上記リスクを軽減するために、

- ・ 事後学習データとする画像に関する詳細なログ（使用環境、診断結果、責任者等）を保存するための機能（6（2）を参照）
- ・ 支援システム等の内部データを利用した自己テスト等により、定期的に性能評価を行い、その時点での性能に使用者に示すことで最低限の性能を確保・確認できる機能
- ・ 上記機能により上記リスクが発生していることが確認された場合に、人工知能を市販時の状態に初期化又は変化直前の状態へと復元する機能
- ・ 上記リスク発生時に、必要に応じて、支援システム等を停止させる機能

等、当該支援システム等の承認申請時に確認された性能を確保するための機能、又は具体的な対策手順をその支援目的に応じて設定しておくことが求められる。

使用者は支援システム等の目的に応じた専門知識を有し、かつ、その使用等に関する教育（トレーニング）を受けた医師を始めとする医療従事者に限定する一方、確定診断、最終判断等、診療行為における最終責任者は医師と規定することが求められる。加えて、使用者の誤使用や支援システム等の目的から逸脱した適応外使用を防ぐため、ユーザトレーニング等の適切な対策を講じておくことが求められる。なお、このユーザトレーニングは関連学会等と連携したものが望ましい。

7. 機器の特性に応じた追加留意事項（案）

（1）事後学習による性能変化が診断支援においてハイリスクとなる支援システム等の場合

例えば、病理診断においては、その診断が臨床における最終診断であることから、どのような形式であってもその性能が市販時から変化することは検出率及び正診率低下の可能性を含むこととなり临床上高いリスクを伴うこととなる。このような支援システム等の場合、その品質、安全性及び有効性を十分に確保することから、現時点では、事後学習プロセス、性能変化及びその検証には必ず病理専門医の関与を要求することが望ましい。

（2）支援システム等の対象画像が動画の場合

消化器内視鏡機器や超音波画像診断装置のように観察像が動画となる場合、人工知能に対する学習データはその動画をどのような形で利用したかを明確にする必要がある。よって、動画のフォーマットや取得方法を明示するほか、学習データへの変換方法を明

らかにする必要がある。

また、当該支援システム等においては、動画の撮影条件及び質が撮像者の技能に左右されることから、学習データとして利用可能となる条件を規定するとともに、その妥当性を明らかにする必要がある。

以上

使用者等が性能変化に関与可能な人工知能技術を利用した 医用画像診断支援システムに関する基本的考え方（案）

1. 背景

今回、人工知能技術を利用した医用画像診断支援システム（支援システム等）に関する評価指標案を作成するにあたり、近年の計算機技術の著しい発展に伴い、実現可能な様々な支援システム等を全て評価指標案の対象に含めるかについて、WG で議論が行われた。特に、人工知能の中でも、近年着目されている深層学習を利用した場合の可能性を考慮し、コンピュータ技術が発達することにより、現在の大型コンピュータと同レベルのスペックの汎用 PC が使用可能となることで、ネットワークを介さずに PC 中に人工知能を有し、特定の臨床現場で事後学習を行い性能が変化するケースを対象に含めるかも議論された。

このケースには、事後学習により生じる性能変化に伴う品質管理を製造販売業者が行うことが難しい支援システム等、例えば、施設又は使用者ごとの事後学習による性能変化が可能なシステム、教師なしデータによる学習が可能なシステム等が含まれる。これらの支援システム等は、解決すべき事項が数多く存在することから、本評価指標の対象外とすることとした。しかしながら、人工知能技術の著しい発展に伴いこれらの支援システム等が近い将来開発される可能性に鑑み、それらの品質、安全性及び有効性を確保するための考え方を別添として取りまとめた。

2. 問題点と解決策について

（1）性能変化

例えば、使用者による事後学習により性能が変化する支援システム等の場合、開発者以外が主体となって事後学習を行うことになるため、製造販売業者の開発意図と大きく異なる方向にその性能が変化・低下し、臨床上許容される水準を下回る可能性が否定できない。このような問題を防ぐためには、適切な質と量のテストデータによる検証に加え、当該支援システム等のリスクを勘案した対策を講じることが求められる。例えば、製造販売業者による品質管理を徹底するための方策を講じる、製造販売業者と同等の品質マネジメントシステムを臨床現場に求める、使用者を対象としたトレーニングを実施することに加え、性能が下限以下となった場合にはそのリスクに応じた追加対策を施すこと等が求められる。

支援システム等の目的は「診断支援」であることから、確定診断時における医師の最終的な判断によりそのリスクを回避することができると考えられるが、可能であれば生じたリスクを漏れなく製造販売業者に連絡する機能、又は手順を必要に応じて設けるこ

とが望ましい。

(2) 責任の所在

その機器の使用に伴う医療行為、適応外使用、不具合等に対する報告に関しては医師が責任を負うことが原則であるが、その原則を一様に適用できない可能性があることに留意して、製造販売業者及び医師の責任範囲を設定する必要がある。

3. 評価に際して考慮すべき事項（品質、安全性及び有効性の評価）

使用者が施設ごとに学習を行うシステムと、自己学習を行うシステムに関する留意点は以下のとおりである。

(1) 使用者による施設ごとの事後学習が可能な支援システム等

特定の医療機関で得られた臨床データや診断支援の結果を元にした事後学習を使用者が実施するケースが該当する。この場合、事後学習データの偏りにより、医療機関ごとに異なる方向に性能が変化し、製造販売業者による品質管理が困難となる。加えて、使用者が事後学習を実施するため、製造販売業者の開発意図と大きく異なる方向にその性能が変化するリスクがある。そのため、性能評価においては、評価指標に記載した事項に加え、製造販売業者は、

- ・ 事後学習プロトコル。特に使用者が施設ごとに行う事後学習データの収集方法や品質管理等も含めたプロトコルの詳細
- ・ 事後学習後に変化したシステムの性能の検証プロトコルに関する詳細。承認申請時のテストデータに対する性能評価結果のみならず、新しいテストデータを追加する場合には、その収集方法や品質の評価法と結果、及びシステムの性能評価結果に関する詳細

に留意し、製造販売業者が事後学習を実施した場合と同等の品質マネジメントを臨床現場に求めるための工夫が必要である。

また、事後学習に加え、変化した性能の評価や、その変化に伴う支援システム等の安全性及び品質の確保を使用者等が行う必要があるため、それらに係る評価は製造販売業者が実施する場合と同等の評価を行う必要がある。そのため、当該支援システム等を有する医療機関内に、以下の要件を充足した責任者の指名と、その責務、性能評価及び品質管理に関する手順書の整備等を求めることになると考えられる。

- ・ 対象となる支援システム等の原理及び学習方法を熟知していること
- ・ 事後学習データの妥当性を示すことができること

なお、その性能変化の範囲が承認申請時に製造販売業者により規定された範囲内であることが前提となるため、性能評価に関しては製造販売業者が承認申請時に規定した検

証方法を用いる必要があり、性能変化が製造販売業者により規定された範囲を超えることが生じないように、評価指標6（4）項「リスクマネジメント」に記載した事項に加え、当該支援システム等にはそれを防止するための機構、具体的には、そのリスクを製造販売業者に連絡する機能等を設けることが強く求められる。

（２）「教師なし」データによる自己学習が可能な支援システム等

2017年10月、Googleにより囲碁用人工知能において「教師なし」データを用いた自己学習により連続的な性能変化が可能となったことがNatureにて報告された。将来、同様の人工知能を利用して医師の最終判断や製造販売業者の介入を要することなく、医療現場での使用中、自律的にその性能を変化させていく支援システム等が出現する可能性は否定できない。加えて、その人工知能がネットワークから切り離された状態で搭載され、製造販売業者が関与しない状態で納入先の医療機関ごとに性能が変化する支援システム等の開発も予想される。しかしながら、このような支援システム等がどのように事後学習を進めるかについては不明な点も多い。

上記のような自律性の高いシステムは、現在の科学技術水準では期待されるレベルの診断支援能力を持つには至っていないが、将来実現した場合には、再学習時に技術者や医師による学習データの作成が不要となり、教師あり学習の場合と比べて支援システム等の自律度は高くなる。また、事後学習用データの収集や事後学習のタイミングも全て自ら判断して実行し、技術者や使用者による介入が完全に不要な自律した支援システム等が登場する可能性もある。このような人による介入度が低い、又は不要な支援システム等では特に品質管理が重要となり、承認申請時の性能を担保していることの確認やリスクマネジメントに関する要求は自律度の上昇に伴って高くなる。これらの要求等を充足するためには、前項までに挙げた各種の要求事項をより高度なレベルで満たすことが求められる。充足すべき事項及びそれらのレベルに関しては現時点で評価指標として具体的に記載することは困難であることから、承認申請に先立って独立行政法人医薬品医療機器総合機構と詳細な相談をしておくことが必要になると考えられる。

以上

IV. 委員報告

- 1：人工知能技術の医療応用：医療情報の観点から
- 2：新 AI 時代におけるコンピュータ支援診断（CAD）の新展開
- 3：深層学習などの機械学習における課題
- 4：病理診断における人工知能による病理診断支援システムについて
- 5：消化器内視鏡領域における AI 技術応用
- 6：超音波検査における人工知能の活用について
- 7：放射線診断分野における人工知能

＜人工知能技術の医療応用：医療情報の観点から＞

香川大学医学部附属病院 医療情報部

横井 英人

1. 背景

日本の医療情報システムは、国民皆保険による統一的な診療報酬制度に基づく医事会計システムが 1970 年代の導入されたことに端を発し、臨床検査、薬剤といった部門システムの開発を経て、2000 年前後から画像システムや、医師や看護師の日常の記録を扱ういわゆる電子カルテの登場を待ち、一定の完成を見ている。

2. 保存されているデータの現状

日常の臨床データがシステムに保存されているのであれば、それを臨床研究などに利用したいと考えることは自然であったが、複数施設のデータ統合は、想像以上に困難が伴うことが近年の様々な臨床研究プロジェクトで明らかとなった。主たる理由は、データ標準化の遅れである。システム上のデータを統合しようとするれば、データ項目名の標準化、および各項目に入力されたデータの表記方法や用語選択の統一性が必要となる。

欧米では、SNOMED-CT と呼ばれる、臨床用語の標準用語集があり、各施設がこの用語集を直接利用していなかったとしても、この用語集へのマッピングを通して、データの標準化を図れる可能性が残されている。しかし日本では、いくつかの分野での標準用語集やコード集が規定されているが、電子カルテへの採用は十分でない。データ項目保存方法の標準化は、厚生労働省の標準規格として規定されている SS-MIX 標準ストレージである程度なされたが、データ収集に十分な項目内容を持つのは、病名、臨床検査（血液・尿検査など）、処方内容の他、診療情報提供書などの分野にとどまっている。

画像分野、特に放射線画像は、DICOM という統一形式が国際的に普及しており、国内でもほとんどのシステムは同規格の基本部分に準拠している。国内の施設は既にほとんどが放射線分野において電子化・デジタル化がなされており、診断はディスプレイ上でなされ、データはシステムに保存されている。したがって、放射線画像データについては短期間に大量の情報を収集し、処理することが可能と考えられる。

3. データ収集の手続き

個人情報保護法の改正・次世代医療基盤法の施行によって、個人情報の二次利用が活性化されると考えられる。しかし、その運用手法については明らかでない部分が多い。また、日本医療研究開発機構 (AMED) の「臨床 ICT 基盤構築・人工知能実装研究事業」などにより、医療情報システムからのデータ収集に関する様々な研究事業が実施されているものの、現時点で、それぞれの事業での匿名化手法やデータ収集手法は統一されておらず、特に医療機関の情報管理手法も定まっていない。大量データの二次利用となれば、オンラインでデータ送受信を行う可能性が高く、その場合には現行のシステムを改修しなくてはならず、またそのデータフローについて継続的な管理が必要になり、システムインパクト・業務インパクトが相当に大きいことが指摘されている。更に手法が統一されていないことで、医療機関はそれぞれの事業に個別な対応を強いられ、その管理業務に関する負担は無視できないものとなっている。

4. AI の学習と評価

ただし、撮像条件なども含めたデータの信頼性を担保するには、データ記録ポリシーの統一化が必要であり、前述のような標準化についても施設間で同じように行っている必要がある。実際に各施設での保存状態を見ると、そこには相当に散らばりがあることがわかる。

人工知能の学習に用いられる画像には、一定量の教師データが用いられると思われるが、そのための情報としての画像診断名・臨床診断名には前述したように診断手法・ポリシーなどが統一されていないので、単純にシステムから取り出すのみで十分かは疑問があり、何らかのバリデーションや、改めて評価者を設定しての再ラベリングや施設からの追加データ収集が必要となり得る。この場合、どのような対応が必要かは、AI の使用目的を鑑みて決定しなくてはならない。その目的が、粒度の粗い分類であり、施設間差などの考慮をさほど要しないのであれば、バリデーションや再ラベリングに重い作業を行わなくてもよいかも知れないし、精緻な推論結果を要求するのであれば、その目的に合致した作業が必要になるであろう。

<新 AI 時代におけるコンピュータ支援診断 (CAD) の新展開>

岐阜大学工学部電気電子・情報工学科／医用画像情報学会
(大学院医学系研究科知能イメージ情報分野 併任)

藤田 広志

1. はじめに

いま AI の第 3 次ブームが到来し、医療分野、とりわけ医用画像診断領域における AI への期待が大きい。しかしながら、医師にとっては、「近い将来、医師の仕事が AI に取って代わられてしまうのでは」との不安の声も聞こえてくる。実際、ディープラーニング (深層学習) で一躍有名になったトロント大学 (現 Google) の Hinton は、「It is just completely obvious that within five years, deep learning is going to do better than radiologists」と、2016 年の講演で放射線科医の読影作業に対して述べている¹⁾。

2016 年末にシカゴで開催された、世界最大規模 (5 万人を超える参加者総数) の北米医学放射線学会 (RSNA2016) でも、AI に関する大々的な各種イベントが実施され^{2), 3)}、引き続き RSNA2017 においてもさらにエスカレートした感がある⁴⁾。このような現状で、AI の影響を受けやすい医用画像を対象とする「コンピュータ支援診断 (computer-aided diagnosis, 以下 CAD)」領域への普及・影響の現状はどうであろうか。本稿では、「新 AI 時代におけるコンピュータ支援診断 (CAD) の新展開」と題して概説する (文献 5, 6 の拙文に基づいている)。

2. これまでの CAD

CAD の研究開発の歴史は古く^{7), 8)}、AI の歴史も概観しながらまとめると、以下のようになる。

1) 黎明期 (1960 年代~1970 年代)

デジタルコンピュータの発明は 1940 年代である。その後およそ 20 年を経て、1960 年の Lusted 論文のコンピュータによる画像解析の必要性や胸部 X 線写真における正常・異常画像の自動分類の提案を始まりとして⁹⁾、CAD の始まりといえる研究が 1960 年代にいろいろ出現している。なお、CAD という用語を用い、CAD として書かれた最初の出版物は、1966 年の Lodwick によるものであったと思われる¹⁰⁾。この論文では、CAD の具体的な開発アプローチを 8 つのステップで示している。ただし、この時代の CAD 研究の多くは、画像の“支援診断”ではなく“自動診断”を目指していた。

AI という言葉が誕生したのは、1956 年に米国のダートマスで開催された会議 (ダートマス会議) においてである。直後の 1958 年には、生物の脳の神経ネットワークをモデルとしたコンピュータ処理の仕組み (ニューラルネットワーク) の基礎となるパーセプトロン (人工ニューロンを 2 層に繋いだ構造) が登場している。もっとも、ニューラルネットワークに関する研究は、すでに 1940 年代から始まっていたが、1960 年代にゲームでの探索による課題解決によって、「第 1 次 AI ブーム」を迎える。しかし、「トイ・プロブレム (おもちゃの問題)」は解けても、現実遭遇する複雑な問題は解けないことが分かり、1970 年代には AI は「冬の時代」を迎えた。

2) 成長期 (1980 年代~1990 年代)

一連の CAD の研究は、その後も疾患の対象領域が広がりつつさらに続けられた^{1 1)}。そして、支援診断を全面に打ち出して CAD という概念・発想でシステムを開発する研究が、1980 年代前半にシカゴ大学の Doi らにより本格的に始まった^{1 2)}。

AI 領域では、この 1980 年代は、コンピュータに「知識」を入れて賢くしようという時代であり、エキスパートシステムとして開発され、「第 2 次 AI ブーム」が起きている。中でも、1970 年代初めにスタンフォード大学で開発されたマイシン (Mycin) が有名であり、これは伝染性の血液疾患を診断し、抗生物質を推奨することができた。

1986 年には、階層構造のニューラルネットワークの学習法としてバックプロパゲーション (誤差逆伝播法) が提案され、ニューラルネットワークは、「学習するコンピュータ」として大きな話題を呼んだ。また、この頃に、福島らがネオコグニトロンという生物の視覚神経路を模倣したニューラルネットワークを発表しており、これはディープラーニングの元祖となる。そして、1990 年頃から、多くの CAD システムに、この 3 層構造のニューラルネットワークが取り入れられ、性能向上に一役を担っている^{1 3)}。

しかし、このような第 2 次ブームは、知識を記述し管理することの難しさが次第に明らかになると、1995 年頃から、また「冬の時代」を迎えてしまう。

3) 実用期 (1998 年 : CAD 元年~2010 年代前半)

1998 年は「CAD 元年」の年であるといわれる。その理由は、米国のベンチャー企業 R2 Technology 社 (現ホロジック社) の開発した検診マンモグラフィ (乳房 X 線写真) 専用の CAD システム「ImageChecker System」が、米国の FDA (食品医薬品局) の認可をこの年に得ており、米国内で商品として販売することに成功したからである (世界初の商用 CAD システムの実現)^{1 4)}。また、米国では、マンモグラフィ CAD の利用に対して、2001 年 4 月から保険の適用が可能になり、CAD の普及に拍車がかかる大きな要因となった。その後、マンモグラフィ (乳がん検出) 以外にも、乳房超音波画像 (乳がん検出)、胸部 X 線写真や CT 画像 (肺がん検出)、大腸 CT 画像 (大腸ポリープ検出) などのいくつかの画像診断領域の CAD も、順次、商用化に成功し、現在に至っている。

これらの CAD の利用方法の定義は、まず、① 医師が画像を CAD なしで最初に読影し、その後、② コンピュータの解析結果を「第二の意見」として利用するものであり、最終診断は必ず医師が行うことと厳格に決められており、「Second Reader」型 CAD と呼ぶ。

これらはすべて CAD の「第 1 世代機」と呼べるが、マンモグラフィ CAD 以外の商用機の普及は、期待に反して進展していない。特に、本邦では、薬機法の承認を得た CAD システムは、マンモグラフィのみに留まっている。

AI の観点からは、この時代には、ニューラルネットワークの限界も分かってきており、次にはサポートベクターマシン (SVM) やランダムフォレストなどの新しい機械学習 (コンピュータのプログラム自身が学習する仕組み) の方法も出現し、CAD の開発に利用されるようになっていく。

3. 第 3 次 AI ブーム時代の CAD

現在は、コンピュータが自律的に学習できるようになったことにより、「第 3 次 AI ブーム」の時代へと突入している。停滞していた AI 研究の分野に、ビッグデータの時代に広がった機械学習と、機械学習の一種であるディープラーニング (ニューラルネットワークの新しい機械学習の方法) の 2 つの大波が襲う。

これまで、画像の中の認識対象の特徴量を、設計者（人間）が苦労して考案・作成してシステムを開発してきたのに対して、ディープラーニングの利点は自ら特徴量を作り出す（すなわち学習する）ことができる点にある（図 1）。ディープラーニングは、層が4層以上のニューラルネットワークの総称である。特に、畳み込みニューラルネットワーク（convolutional neural network: CNN）と呼ばれるディープラーニングが良く使われている。



(a) 従来型の機械学習によるCAD



(b) ディープラーニングによるCAD

図 1 従来型学習とディープラーニングによる学習の比較

このような中、前述の RSNA2016 や RSNA2017 では、急激な AI ブーム、ディープラーニングブームが見られ²⁻⁴⁾、CAD システムへの AI 技術（特にディープラーニング）応用の研究報告がこれまでに比べ大幅増となった。CAD の一部の機能にディープラーニングを利用して（例えば、偽陽性候補の削除処理）検出、あるいは分類の性能を改善しているものもあれば、ある特定の領域（例えば、胸椎・腰椎、乳房）を領域分割（セグメンテーション）するツールとして利用するものもあった。なお、医用画像は、一般画像のように正解ラベル付きの大量のデータを収集することは容易ではないため、一般画像でディープラーニングを学習したものを目的の医用画像に適用する転移学習（transfer learning）を利用したものも多い¹⁵⁻¹⁷⁾。また、画像をディープラーニングで取り扱うのみならず、自動で読影レポートも作成する機能をも有するシステムもあり、これは興味深い。

さらには「完全ディープラーニング型」と呼べる、ディープラーニングへの画像入力とその病変の有無のみの情報付与で、システムを構築しようとするタイプの CAD も見られた。特に、ある企業では、その目指すところは、最初にコンピュータが病変の検出処理を行い（“First Reader”）、医師が読影すべき異常がありそうな画像をふるいわけし、これにより医師の負担を減らそうというものであるという。この CAD 読影方法は、現状の「すべての画像を医師は診（見）なければならない」という原則を外れることになる。果たして、それが可能なだけの性能を出すことができるのかどうか、今後の展開が楽しみである。

新しいタイプの CAD として注目できるものとしては、QView Medical 社が開発した全乳房超音波画像（3D automated breast ultrasound: ABUS）のための CAD システム「QVCAD」があり、2016 年 11 月、FDA の承認（premarket approval: PMA）を得ている。

特徴は、“Second Reader”ではなく“Concurrent（同時）Reader”で、“同時読影型の

CAD”として、初めて FDA の認可を得ている点である。ここでもディープラーニング技術が使われている。また、同様に多量の画像の読影が迫られる乳房トモシンセシス画像を対象とした CAD も、2017 年 3 月に iCAD 社の CAD システムが FDA の承認を得ている。同時 CAD は、CAD の「第 2 世代機」ともいえる。よって、もし上記の”First Reader”の CAD が出現すれば、これはスクリーニング現場で、真に期待されている読影を軽減できるものであり、このようなシステムは、CAD の「第 3 世代機」と呼べるであろう。

これらの CAD は、すべてコンピュータ支援検出型の CAD (computer-aided detection: CAdE) として分類されるものである。一方、検出領域が良性であるのか悪性であるのかなどのさらに一步踏み込んだ画像診断支援は、CADx (computer-aided diagnosis: CADx) として区別されるが、FDA 承認を得た商品はこれまで皆無であった。ところが、2017 年 7 月に、Quantitative Insights, Inc (QI) という米国のベンチャー企業が、乳房画像解析の CADx を開発し、CADx としては初めて FDA 認可に成功している。ここでも、ディープラーニング技術が使われているという。

現状の CAD 開発の流れを、CAD の利用方法に応じて分類すると (進化形態でもある)、図 2 のようになるであろう。まだ商用化は実現していないが、上述の”First Reader”型 CAD、さらには自身でどんどん賢くなる“事後学習機能付きの CAD”が出現するであろうと予想される。その先には、いよいよ“自動診断”が見えてくる。

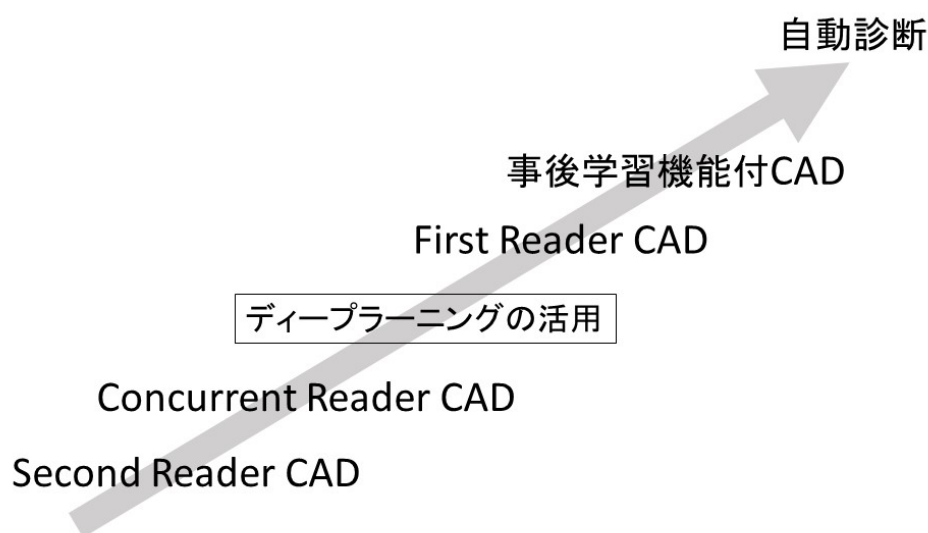


図 2 CAD の利用形態・進化形態による分類

4. 次世代型 CAD の開発に向けて

AI の段階には、レベル 1 から 5 までである^{18, 19)} (図 3)。レベル 1 の AI は、AI 搭載をうたった家電製品で (単純な制御プログラム)、レベル 2 は質問応答システム、お掃除ロボットや一般的な将棋ソフトウェアのレベルである (「知識」を使った AI で、推論・探索が可能)。レベル 3 になると、検索エンジンやビッグデータ分析に活用されるもので、機械学習が行われるようになり、人間から特徴量を教えられて学習する。レベル 4 では、さらに高度な分析

が可能になり、ディープラーニングが取り入れられ、人間が特徴量を教えなくても自力で獲得（学習）する。車の自動運転や、昨今の囲碁 AI などにはこれに該当する。しかし、これらのレベルの AI は、特定の目的に限定されるもので、「特化型 AI」と分類される。そして、レベル5になると、人間のように（あるいはそれ以上に）何でもできる AI で（例：鉄腕アトム！）、「汎用 AI」（artificial general intelligence: AGI）と呼ばれるが、これはまだ実現できていない。レベル5の AI は「強い AI」とも呼ばれ、それ以外は「弱い AI」とも呼ばれることがある。

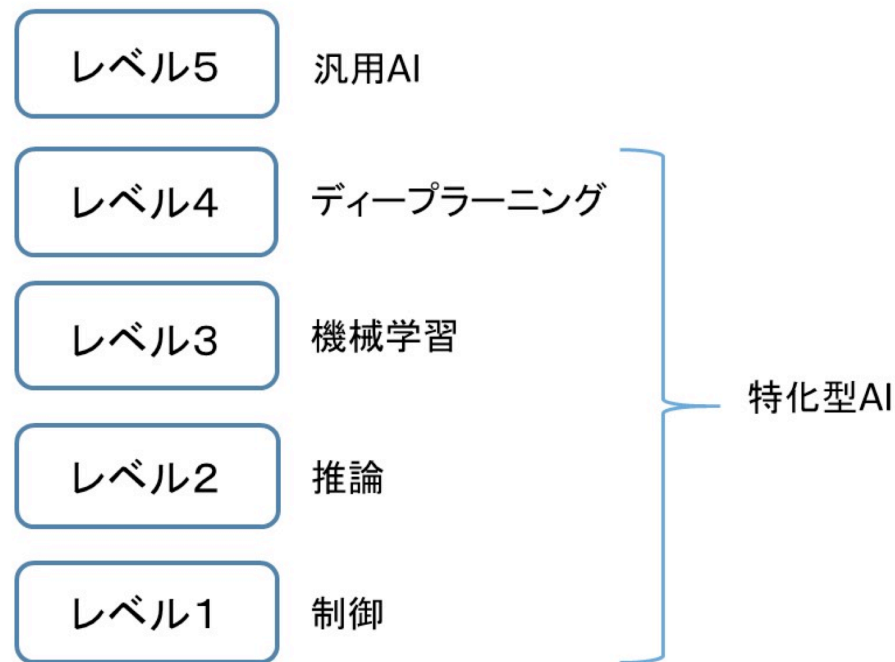


図3 5段階の AI レベル

これまでの商用化されている CAD はすべてレベル3であり、いま開発が進められている（一部はすでに商用化）CAD はレベル4である。

さて、このように新しい AI 技術（特に、ディープラーニング）が CAD に取り入れられてきており、Second Reader から Concurrent Reader、そして、First Reader CAD の方向に向かっている。“事後学習機能付き CAD” も出現すると考えられ、これには本評価指標案でも示唆しているように、解決すべき問題が多く存在する。

CAD を開発する際には、前述のように（図1）、従来型の機械学習では、医師の読影過程をコンピュータのアルゴリズムに書き込んで実行するには、「入力画像＞特徴抽出＞識別処理＞識別結果」のプロセス（手順）に従っている。すなわち、マニュアルによって特徴量を設計するため、相当な時間を要した。また、アルゴリズムが複雑にもなった。一方、機械学習の代表格としてのディープラーニングでは、“十分な質・大量” の入力画像とその正解（病変の有無あるいは位置、良悪性の判定結果、検出対象の輪郭図など）をペアで用意すれば、それだけで良い。

有名な「Google の猫」の実験では、猫の概念をディープラーニングで抽出・表示化して話題になったが、YouTube から取り出した 1000 万枚の画像と、1 万 6000 個ものコンピュータプロセッサをつないだ大規模なディープラーニングを構築している。ここに2点の重要なポイントがある。すなわち、計算機のパワーと学習に使うデータ量である。計算機の問題は、

GPU (graphics processing unit) と呼ばれる専用のプロセッサで、並列演算を高速で行える。また、量子コンピュータの開発などで、今後もどんどん速くなるので、解決されるであろう。しかし、data-hungry としばしば呼ばれるディープラーニングを学習するための大量の画像データを収集することは、医学領域では必ずしもそれほど簡単な話ではない。そのような中でも、最近、一般画像に比べると大規模とはまだ呼べないが、10 万枚を超える規模の医用画像データを利用した、ディープラーニングによる成果を出している研究事例が出てきている (例えば、眼底画像²⁰⁾、胸部X線画像²¹⁾、皮膚がん検査画像²²⁾ において)。個人情報保護の観点からの問題があり、そもそも病変の画像データは「スモールデータ」である。よって、この問題解決には、国家レベルのプロジェクトでデータ収集を行うなどの枠組みが必要であり、検討が進められている。また、少数データにも対応できる新しい学習法開発も鋭意、進められている²³⁾。

従来の米国 FDA や本邦の薬機法の CAD 承認では、コンピュータアルゴリズムのパラメータを勝手に変更することは許されず、修正申請する必要がある。すなわち、コンピュータが自己学習して、臨床現場でどんどん賢くなるような CAD の枠組みは現状では許されない、ということになる。AI、特にディープラーニングでは、どのような処理がディープラーニング内で行われているのかは完全にブラックボックスであり、なぜそのような判断が下されたのか、正確には分からない。よって、この対応も重要な検討課題であり、説明責任を伴う AI の研究が進められている (“ホワイトボックス化”)。

5. 参考文献

- 1) <https://www.youtube.com/watch?v=NoIB7DXRwl4>
- 2) 藤田広志：人工知能 (AI) の最新動向, INNEVISION, 32 (2), 34-36 (2017).
- 3) 藤田広志, 木戸尚治, 原 武史: CADシステムの最新動向, INNEVISION, 32 (2), 36-39 (2017).
- 4) 藤田広志, 木戸尚治, 原 武史: CADシステムの最新動向-AIの動向も含めて, INNEVISION, 33 (2), 28-30 (2018).
- 5) 藤田広志: AIがもたらすCADシステムの変革, INNEVISION, 32 (7), 10-13 (2017).
- 6) 藤田広志: AIのコンピュータ支援診断 (CAD) への展開, 「AI導入によるバイオテクノロジーの発展」, 株式会社 シーエムシー出版, 印刷中・2018年2月上旬発行予定.
- 7) 藤田広志: コンピュータ支援診断 (CAD) の技術史, 電気学会誌, 133 (8), 556-559 (2013).
- 8) 藤田広志: 実践 医用画像解析ハンドブック, p.518-p.533, オーム社 (2012).
- 9) L. B. Lusted: Logical analysis in Roentgen diagnosis, *Radiology*, 74 (2), 178-193 (1960).
- 10) G. S. Lodwick: Computer-aided diagnosis in radiology. A research plan, *Investigative Radiology*, 1 (1), 72-80 (1966).
- 11) 鳥脇純一郎, 舘野之男, 飯沼 武 (編): 医用X線像のコンピュータ診断, シュプリンガー・フェアラーク東京 (1994).
- 12) K. Doi: Computer-aided diagnosis in medical imaging: Historical review, current status and future potential, *Computerized Medical Imaging and Graphics*, 31, 198-211 (2007).
- 13) 岡部哲夫, 藤田広志 (編著): 医用画像工学, 医歯薬出版株式会社 (2010).
- 14) 長谷川 玲: 世界で初めて商品化されたマンモグラフィ用CAD-ImageChecker-, 日本放射線技術学会雑誌, 56 (3), 355-358 (2000).

- 15) H. C. Shin, H. R. Roth, M. Gao, *et al.*: Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning, *IEEE Transactions on Medical Imaging*, 35 (5), 1285-1298 (2016).
- 16) B. Q. Huynh, H. Li, M. L. Giger: Digital mammographic tumor classification using transfer learning from deep convolutional neural networks, *Journal of Medical Imaging*, 3 (3), 034501 (2016).
- 17) E. Ribeiro, A. Uhl, G. Wimmer, M. Häfner: Exploring deep learning and transfer learning for colonic polyp classification, *Computational and Mathematical Methods in Medicine*, vol.2016, 6584725 (2016).
- 18) 坂本真樹 : 坂本真樹先生が教える人工知能がほぼほぼわかる本, オーム社, 2017.
- 19) 松尾 豊 : 人工知能は人間を超えるか ディープラーニングの先にあるもの, KADOKAWA, 2015.
- 20) V. Gulshan, L. Peng, M. Coram, *et al.*: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs, *JAMA*, 316 (22), 2402-2410 (2016).
- 21) X. Wang, Y. Peng, L. Lu, *et al.*: ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, <https://arxiv.org/abs/1705.02315v2>
- 22) A. Esteva, B. Kuprel, R. A. Novoa, *et al.*: Dermatologist-level classification of skin cancer with deep neural networks, *Nature*, 542, 115-118 (2017).
- 23) 岡野原大輔, 大田信行 : AIは医療分野にどのような変化をもたらすか, 国際医薬品情報, 1076, 8-15 (2017).

＜深層学習などの機械学習における課題＞

東京農工大学大学院 工学研究院

清水 昭伸

1. 概要

深層学習などの高度な機械学習を診断支援に利用する場合、幾つかの問題が想定される。ここでは、診断支援システムの評価基準作成の際に重要となる可能性の高い点について述べる。

以下ではまず、深層学習特有の性質であるブラックボックス性について取り上げる。次に、深層学習などの高度な機械学習が備えている入出力の強い非線形性や、出荷後の継続的な性能変化について触れる。また、学習やテスト用のデータベースに関する最近の特徴について幾つか指摘する。最後に、深層学習の将来とその他の課題について述べる。

2. ブラックボックス性

この性質は、ニューラルネットワークに特有である。学習済みのネットワークの内部はブラックボックスであり、その動作原理は人が見ても分からないことが多い。そのため、例えばある診断が導かれたとしても、その診断の根拠の説明が困難である。以下では、このようなシステムに対して想定される幾つかの問題と対策について述べる。

1) トラブルが生じた際の原因の切り分けや改良が困難

例えば、明らかな異常であるにもかかわらず正常であると診断するトラブルが生じた場合、ブラックボックスであるために原因の切り分けができない。そのため、ネットワークのどの部分を改良すれば良いかや、どのように再学習をすれば良いかが明瞭ではない。

2) 情報セキュリティの脆弱性

ハッキングにより、ネットワークの一部あるいは全体が書き換えられ、誤った診断結果が出力される可能性がある。ブラックボックスの場合には、このハッキングによる書き換えに気づきにくい。

3) 対策例

判断の根拠について説明できる人工知能（Explainable AI; XAI）の研究が進められている¹⁾（図 1 参照）。この図では、ブラックボックスからの出力に対して、根拠を示す画像（左中段の heatmap）を提示しており、雄鶏と判定した根拠が鶏冠の付近に存在することが分かる。現時点では複雑な認識問題に対して人と同程度の説明ができるレベルには無いが、将来は、このような技術の進化によってブラックボックスも説明できるようになり。上で指摘した問題が緩和、あるいは解消される可能性もある。

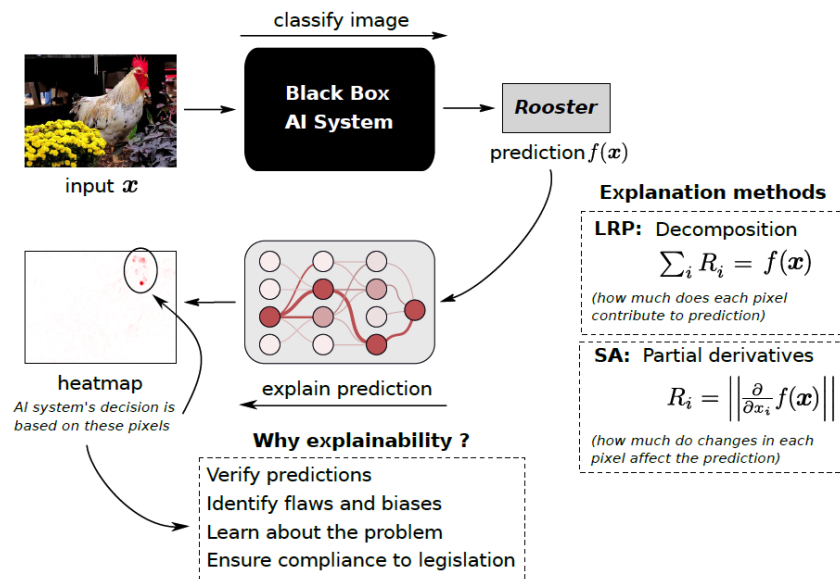


図1 説明可能な AI (XAI) の例 (文献1の Fig.1 から転載)

3. 入出力の非線形性

入力に対して出力が非線形に変化する性質は、深層学習に限らず、高度な機械学習技術、例えば、非線形カーネルを用いたサポートベクターマシンなどが備えている性質である。最近の層数が多い深層ネットワークの場合には、特にこの非線形性が強くなり、入力のわずかな変化に対して出力結果が大きく変わることがある。この性質によって引き起こされる問題として、文献2の敵対的な例 (adversarial example) が良く知られている。図2に一例を示した。これは、パンダの画像に小さなノイズを加えた結果、人にとっては同じパンダにしか見えないが、深層ネットワークは高い信頼度でテナガザルであると認識した例である。入出力の非線形性は、複雑なパターン認識問題を解くためには重要な性質であり、それ自身は悪いことではない。しかし、図2は、わずかなノイズによって誤った判断をする可能性があることを示唆している。深層ネットワークのように非線形性が非常に強い技術は、未知データに対する振る舞いが予想困難である場合も多く、予想もしない誤りが発生する可能性があることに注意をする必要がある。

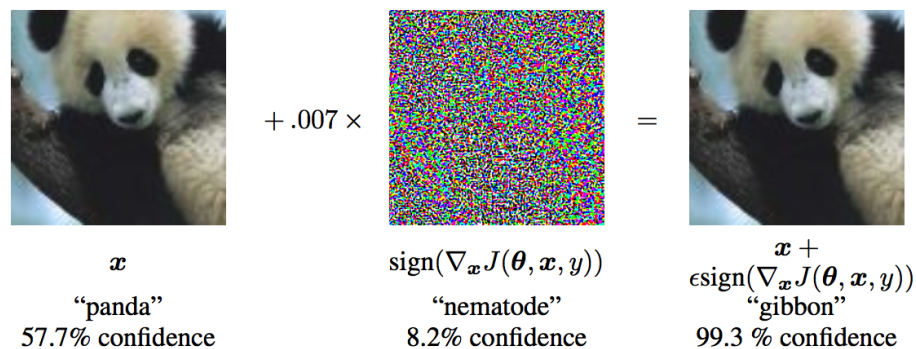


図2 深層ネットワークに対する敵対的な例 (文献2の Fig.1 から転載)

4. 出荷後の継続的な性能変化

ビッグデータと分散並列計算環境の整備に支えられ、深層学習などの機械学習のアルゴリズムは、技術的には任意の時刻において再学習が可能となった。出荷後の再学習によって、システムの性能は出荷前の状態から変化するようになり、臨床現場の実情に合った最適な支援が可能になる。しかしその一方で、性能変化が良くなる方向だけでなくその反対もありうることから、例えば製造販売承認事項との整合性や、リスクマネジメントについて難しい問題が生じる可能性がある。そのため、再学習によって変化する性能の評価・管理法について、科学的な根拠をもって適切に決めておくことが重要となる。また、発生するリスクの大きさは、再学習の主体者、タイミング、学習用やテスト用データの性質などの様々な要因により変わることから、この点も踏まえた議論が必要となる。

5. データベースに関する最近の特徴

医用画像を対象とするシステムの開発において、自然画像による学習を最初に行うことで、深層ネットワークの性能が向上することが知られている。また、将来、Web上に大量に存在する素性が分からないデータを学習に用いることで性能が改善することも想定される。以下では、最近の機械学習に特徴的なこれらの点に注目する。学習、バリデーション、テストなどの様々な目的に使われるデータセットの信頼性の記述や評価の基準の作成の際には、以下の視点が重要となる。

- 1) 転移学習と呼ばれる学習方法の場合、本来の対象とは異なる自然画像を用いた学習を最初に行うことがある。その延長上には、Web上の大量の自然画像や、素性（撮影者、撮影日付、撮影条件、撮影装置など）が良く分からない大量の医用画像を利用することも想定される。素性の明らかな医用データしか用いてこなかった医療機器開発の従来の方法論が、変化しつつある。
- 2) 教師あり学習には、入力画像に対する教師としてのラベルやカテゴリ名が必要であり、一般にこれらの信頼性の評価は重要である。例えば、複数の専門家が合議で作成したアノテーション画像や、病情報に基づくラベルは信頼性が高く、そうでないものは低く、学習結果へ悪影響があると考えられてきた。しかし、最近よく用いられるようになった半教師あり学習や Weak Label が利用可能な学習アルゴリズムでは、一部の教師ラベルのみの場合や、教師ラベルが不正確な場合も扱えるようになった。教師無し学習の様に、ラベル無しで学習が行えるものも登場している。

このようなアルゴリズムが登場した背景には、膨大な画像データの蓄積がある。例えば、カプセル内視鏡や大腸内視鏡のように大量の画像を用いて学習する場合、すべてに対して病理学的に正確なラベルをつけることはほぼ不可能に近く、仮に正確なラベルを付与できてもそれに見合うだけの性能の改善が期待できないケースもある。一般には、ラベルの信頼性に関する正確な記述は重要であり、教師データの信頼性が不十分であることの影響や、学習後の結果が受容可能であることの検証が必要であるが、常にラベルが正確であることが必要か否かは問題依存である。したがって、教師ラベルなどの信頼性の評価を、問題によらず一様に厳しく求めることは適切ではないとも考えられる。

（注：テストデータのラベルについては従来と同様に高い信頼性が求められる）

- 3) 学習データを補うために、初期の学習データに対して線形変形（回転、拡大縮小、平行

移動など)や非線形変形(Free form deformation や Radial Basis Function による変形)を適用したり、生成モデル(例:自己符号化器など)を用いて人工的にデータを生成することがある。Data Augmentation と呼ばれる技術であり、実際の画像に加えて学習させることで性能が改善する。この技術では、変形や人工データ生成のためのアルゴリズムがポイントとなる。例えば、適用する変形が想定されるデータの分布と整合していれば効果が期待できるが、不適切な場合には悪影響も想定される。

- 4) 診断支援等のシステムの性能を評価するテストデータは、学習に用いないことはもちろん、ハイパーパラメータ(ネットワークの層数や学習回数など)を決めるためのバリデーション用データとも区別して管理する必要がある。その上で、以下のように、テストデータを用いて学習後のシステムを繰り返し評価する過程で、テストデータの信頼性が失われてしまう例があることにも注意しなければならない。例えば、10通りの学習を行い、学習済みの10通りのシステムをテストした後、その中からもっとも性能が良いシステムを選択して運用した場合である。一見、テストの目的にしか使っていないようにも思えるが、システムを選択も設計に含まれる重要なステップであり、上記の例ではそこにテストデータを利用している。この場合のテスト結果の信頼性は低くなる。

以上のことから、理想的には、テストデータの管理とそれを用いた性能評価は、第三者機関が実施することが望ましい。例えば、再学習後のシステムの性能が基準を満たしていることの確認は第三者機関が実施し、性能管理に利用することが望ましい。しかし、実際にはコスト面などで困難なことも多い。従って、開発者が、上記の例などに配慮しながらテストデータを適切に管理することが重要となる。

- 5) 大規模な公開データベースをテストなどに利用することは、医療機器の開発において非常に有益である。例えば、米国 MIT で開発された心電図データベースは有名であり(MIT-BIH Arrhythmia Database)³⁾、このデータベースを使うことで心電計などの開発・普及が加速化された事実がある。今後は、大規模な公開画像データベースの整備を一層加速させることが重要となる。

6. 深層学習の将来とその他の課題

深層学習が様々な認識システムの性能を大きく向上させたことは事実である。しかし、このまま技術的特異点⁴⁾に結びつくほどの実力があるか否かは現時点では不明である。深層学習の発展を支えてきた技術は主として発見的なものであり、理論的な解明が不十分であることが、将来の見通しを不透明にしている。深層化による性能向上も一段落し、単純に層数を増やしても思うようには性能が伸びず、過学習により性能が低下する例も報告されている^{5, 6)}。

一方、今回のブームの背景にあるビッグデータと分散並列計算環境に関する研究は、当面は発展を続けると予想される。この点は、特に多くのデータと計算資源が必要な深層学習にとっては、将来を予想する上でプラスの材料の一つである。そこに理論的な解明が加わることで、本質的なブレークスルーにつながる可能性もある。

7. 参考文献

- 1) W. Samek, T. Wiegand, K. R. Mueller: Explainable Artificial Intelligence: Understanding Visualizing and Interpreting Deep Learning Models, arXiv:1708.08296v1 [cs.AI] 28 Aug 2017.
- 2) I. J. Goodfellow, J. Shlens, C. Szegedy: Explaining and Harnessing Adversarial Examples. ICLR 2015.
- 3) G. B. Moody, R. G. Mark: The impact of the MIT-BIH Arrhythmia Database. *IEEE Eng in Med and Biol* 20(3):45-50 (May-June 2001). (PMID: 11446209).
- 4) R. Kurzweil: The Singularity Is Near: When Humans Transcend Biology 1st Edition. The Viking Press; 1st edition (September 22, 2005).
- 5) K. He, X. Zhang, S. Ren, J. Sun: Deep Residual Learning for Image Recognition, *CVPR*, pp.770-778, 2016.
- 6) O. Kuchaiev, B. Ginsburg: Training Deep AutoEncoders for Collaborative Filtering, arXiv:1708.01715, 2017.

<病理診断における人工知能による病理診断支援システムについて>

東京大学医学部 人体病理学・病理診断学分野
地域連携推進・遠隔病理診断センター センター長／日本病理学会
佐々木 毅

1. はじめに

病理診断は、疾患の「最終診断」であり、その診断が患者治療にそのまま反映されることが多い。そのため人工知能（深層学習等）による病理診断支援システム（以下支援システム）は、その精度管理が重要であるとともに、どのような形式であってもその性能が市販時から変化することは、検出率、正診率の低下につながる可能性があり、臨床上高いリスクを伴うこととなる。

したがって品質、安全性及び性能を十分に確保する必要があり、この性能の確保のために、事後学習プロセス、性能変化及びその検証を常に行うことが重要となる。そのため支援システムの使用には病理専門医が関与することが望ましく、十分に配慮して使用することが重要である。

2. 人工知能による病理診断支援システム使用時の留意事項

1) HE 染色標本・病理デジタル画像について

通常の病理診断に使用するヘマトキシリン・エオジン染色標本（以下 HE 染色標本）に関しては、現在、染色工程等が標準化されておらず、施設ごとに染色性に違いがあることが指摘されており、支援システムの診断精度に影響を与えることが報告されている¹⁾。

さらに支援システムでの病理診断の際には、通常病理デジタル画像が使用される。この病理デジタル画像はバーチャルスライドスキャナーという特殊装置で作製されるが、病理デジタル画像には放射線画像での DICOM のような共通フォーマット形式がなく、各スキャナーベンダーによってその形式が異なっている。このためデジタル画像およびモニターに再現する際のビューワの特性が各ベンダーによって異なっており、承認された支援システムを開発した際に用いられたスキャナーとは必ずしも一致しないことが予想される。

したがってオンプロミスおよびクラウド等で支援システムの使用を開始する際には、開発メーカーの立会いの下、各施設の HE 染色標本を各施設のスキャナーでデジタル化した病理デジタル画像を用いて、メーカーが提唱する性能が十分に出ているか等の検証を行った後に運用を開始することが望ましい。承認時の性能が十分に出ていない場合には、メーカー立会いの下、各施設での調整が必要である。

2) 病理診断結果の解釈に関して

支援システムによる病理診断結果に関しては、最終的には病理専門医による結果の確認が必要である。厚生労働省で開催された「保健医療分野における AI 活用推進懇談会」の報告書²⁾にもあるように、最終診断は医師が行うこととしている。特に支援システムによる病理診断においては、最終診断のチェックアウトは、通常の病理診断同様、病理専門医による診断の最終確認が必須である。

3) 支援システム性能変化への対応

先述のように人工知能（深層学習等）の特性上、事後学習による診断アルゴリズムの変化により、その性能が向上することが期待される反面、意図しない性能変化が生じ、承認時の性能を下回る可能性も示唆されている。このような利点とリスクとを勘案し、支援システムには、継続的な性能検証によるほか、学習用データの要件や学習のプロセス等を規定し、学習の品質を確保するための対応策が求められることになる。日々の使用に伴って事後学習が行われ、性能が連続的又は相当程度高い頻度で変化する場合には、特にその内容、プロセス管理に加え、変化した性能の検証及びその管理が重要となる。そのため、支援システムの品質、安全性及び有効性を確保することが必要であり、各施設で病理標本デジタル画像のテストセットを準備するなどして、病理専門医による定期的なチェック、キャリブレーションを行うことが重要であり、適用可能な検証方法、変化した性能が規定された範囲を逸脱した際の対応策等を定めておくことが求められる。

3. まとめ

人工知能による病理診断支援システムは、今後、日本における慢性的な病理医不足の解消に貢献すると同時に、多忙な病理医の支援、時間創出に寄与することが期待される。現在、病理医不在病院を受診した際に病理診断が享受できないという、受診する医療機関によっておこっている医療の不平等を解消するためにも、人工知能による病理診断支援システムは今後非常に有用なツールとなることが期待される。しかしながら、診断精度の確保・管理、および性能変化等には十分な配慮が必要であり、支援システムを活用する病理専門医の人工知能に関する教育、人材育成が必須であると考ええる。

4. 参考文献

- 1) Y. Liu, K. Gadepalli, M. Norouzi, G. E. Dahl, T. Kohlberger, A. Boyko¹, S. Venugopalan, A. Timofeev, P. Q. Nelson, G. S. Corrado¹, Jason D. Hipp, Lily Peng, and Martin C. Stumpe. Detecting Cancer Metastases on Gigapixel Pathology Images. *Google Brain*, 7(3), 2017.
- 2) 保健医療分野におけるA I活用推進懇談会 報告書
(<http://www.mhlw.go.jp/file/05-Shingikai-10601000-Daijinkanboukouseikagakuka-Kouseikagakuka/0000169230.pdf>)

< 消化器内視鏡領域における AI 技術応用 >

京都第二赤十字病院／日本消化器内視鏡学会

田中 聖人

1. 概要

消化器内視鏡、いわゆる胃カメラや大腸カメラと呼ばれる診療のニーズは年々増加しており、健康で生活できる社会の実現のためには症例数の多い、消化器領域の病変の早期発見は今後も重要度が増すと思われる。しかしながら消化器内視鏡医のマンパワーには限りがあるとともに、今後胃を扱う上部消化管内視鏡から、大腸領域を扱う大腸内視鏡に中心が移っていくことを考えるに、内視鏡医の数に限界があることはすなわち、内視鏡診療が生む国民への利益を損なうことになりかねない。AI による内視鏡診断が可能になることで、見落としの少ない良質な診療が容易に行いうるようになることは重要なことであると思われる。

また消化器内視鏡分野は世界でもわが国の優位性が非常に高い。機器メーカーも世界の市エアのほとんどを日本メーカーがカバーしているとともに、消化管内視鏡の分野においては世界をリードする先進的な研究、実地臨床が行われている。この意味においてもわが国において消化器内視鏡領域で AI 研究が進むのは必須と言ってよいであろう。

2. 原理、構成等

消化器内視鏡領域の AI は大きく下記の二つに分類されると思われる。

1) 機器搭載型

内視鏡施行時に、実際の病変に対し病変の発見と見落としの回避ができるもの、あるいは診断の補助となるもので、処理はリアルタイムに行われることになる。このタイプは内視鏡機器や Scope に何らかの機能が搭載されることが条件となり、勢い開発主体は機器メーカーとなり、AI が利用可能な特殊な機器を用いて行うことになる。また AI 解析の処理スピードに依存してしまうため、ある程度の技量が術者に求められる。

2) バッチ処理型

一定の複数の画像を撮影し、これらを一括で処理するものとなる。これは内視鏡機器そのものに搭載する必要はなく、画像ファイリング装置に集積した画像群を解析し、異常所見などを指摘するものである。この方式では大量の画像を一括処理するため、検診領域に重要な役割を果たすと思われる。さらにカプセル内視鏡のように、検査終了後に多量の画像が得られるような検査形態にも応用可能である。このような技術が実現することによって、一部の医療でも行われているように、医師以外による検査が可能となり、処理件数が飛躍的に伸びる可能性を秘めている。このタイプのデメリットはリアルタイムに結果が得られないということ、そして AI での判断が可能となるような良好な画像がないと認識率が低下することが予想される。さらに先進的な研究に比べて派手さに欠ける側面があり、技術開発チャレンジにおけるモチベーションの問題がある。

3. 臨床でのニーズ

前述の双方での有用性は目的が異なるものの、大きなメリットをもたらすと思われる。機

器搭載型のものにおいては、検査、治療施行時にリアルタイムの処理が可能であるため、経験の少ない医師、将来的には医師以外の職種が内視鏡検査を行ったとしても、見落としなどのリスクが軽減されることが期待できる。またより高度な鑑別診断に応用することで、人間が行うよりも正確な内視鏡診療が可能になることが予測される。

バッチ処理型のものは、大量の検査画像を自動認識して検討できることから、検査・治療のすそ野が広がることが想定される

4. 現在の研究状況

本領域の AI 研究はわが国を中心に展開されている。消化器内視鏡診療施行時にリアルタイムで自動的に病変を指摘できるもの、癌、非癌の分類を AI で行うものなどがすでに開発されており、実地臨床への適応が待たれる状況である。領域としては大腸内視鏡分野における検討が進んでいる。いずれも現状では、先進的な領域において機器搭載型が散見される。しかしながら臨床への応用はまだ為されていない状況である。

5. 今後の展望及び解決すべき課題等

他領域における AI 解析と同様に正確な病変マーキングの手法と精度が最も大きな問題であると思われる。日常臨床において良質な画像に正確なマーキングが行いうるような方法論の開発が急務である。現在内視鏡診療の領域では画像ファイリング機能を含むデータベースが広く普及しており、AI 専用の外付けのアプリケーションを利用するのではなく、内視鏡データベース側にマーキングを容易に行い、そしてマーキングがなされた画像を簡便に出力する機能を持たせる必要が出てくる。また画像につけるべき文字情報のタグ付けに関しては、内視鏡学会として取り組んでいる Japan Endoscopy Database ; JED プロジェクトで標準化を行い、詳細で正確なタグ付けは可能になると思われる。このマーキングとテキスト情報のタグ付けが確実かつ容易に行い得るようになることで、大量の教師画像の集積が実現できる。

しかしながら学術的な画像を得るためには、現状では患者側へのインフォームドコンセントが必要であり、いわゆる倫理的な問題の解決は必須であるといえる。今後もこの分野では法整備が進むと思われるが、早急な整備が望まれる。

6. 参考文献

- 1) Y. Kominami, S. Yoshida, S. Tanaka, Y. Sanomura, T. Hirakawa, B. Raytchev, T. Tamaki, T. Koide, K. Kaneda, K. Chayama: Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. *Gastrointest Endosc.* 2016;83:643-9.
- 2) Y. Mori, S. Kudo, K. Wakamura, M. Misawa, Y. Ogawa, M. Kutsukawa, T. Kudo, T. Hayashi, H. Miyachi, F. Ishida, H. Inoue: Novel computer-aided diagnostic system for colorectal lesions by using endocytoscopy (with videos). *Gastrointest Endosc* 2015;81:621-9.
- 3) AI in Endoscopy-Deep Learning for Optical Biopsy of Colorectal Polyps in Real-Time on Unaltered Endoscopic Videos, M. F. Byrne, N. Chapados, F. Soudan, C. Oertel, M. L. L. Pérez, R. Kelly, N. Iqbal, F. Chandelier, D. K. Rex, *Gastrointest Endosc* 2017; 85(5 Supplement):AB364-5.

＜超音波検査における人工知能の活用について＞

東京慈恵会医科大学／日本医用画像工学会

中田 典生

超音波検査は、X 線写真・CT・MRI などの放射線科領域の画像診断と同様、技術的には画像診断に分類される検査である。しかしながら、超音波検査は CT や MRI と異なり、検査中、あるいは検査直後にリアルタイムな診断が行われるという特徴があり。この点は消化管内視鏡検査に類似している。また、超音波検査は低侵襲で被曝もないこと、また近年、装置の小型化により携帯可能な装置が開発実用化されて普及したことなどにより、本邦では超音波検査は医師のみならず、臨床検査技師や診療放射線技師によって広く検査が施行されている。他方、超音波検査および診断は操作者の熟練や豊富な経験が必要な検査、すなわち術者依存性のある検査であり、広く普及しているにもかかわらず検査手技や診断に熟練が必要であるという矛盾を抱えている画像診断検査といえる。これらの矛盾を解決すべく自動的ないしは半自動的に超音波画像を取得できる装置も開発されているが、装置が高額になってしまい未だ広く普及しているとは言い難いのが現状である。

このような状況で、人工知能とくにディープラーニングの超音波検査への応用の科学論文は近年増加しており、CT や MRI など放射線科領域とほぼ同等の論文数が発表されている（図 1）。この理由として、超音波がリアルタイムに動画像として記録可能であること、静止画像が被曝を配慮せずに自由に撮像できること、撮像された画像の画素数がマンモグラフィなど通常の X 線写真に比べて低解像度であることなどの理由により比較的多くの教師学習用データを取得できることなどが考えられる。特に、肝臓や骨盤内臓器を含む腹部超音波検査、乳腺・甲状腺、心臓超音波検査の分野での科学論文が多い。また、乳腺・甲状腺の分野では海外で機械学習による診断支援装置として実用化・商業化された製品がすでに存在する。

超音波検査では解剖学的な臓器の判定などの教師学習用データの取得が特に容易であるため、初期のディープラーニングの研究は進んでいる。しかし実用化に向けてディープラーニングを超音波診断に応用する場合は今後以下の点に考慮するべきである。

- 1) 超音波診断は、他の画像診断（例えば CT や MRI）に比べて客観性に乏しく、診断や所見の採用も術者の熟練度に依存している。従って、超音波検査の診断や所見をそのまま教師学習用データに用いる場合は、できるだけ主観的診断・所見を排除した教師学習用データの作成が望ましい。
- 2) 従来の超音波診断や診断ガイドラインは、超音波検査の専門家の総意により作成された診断学の集大成である。しかしながら、ディープラーニングによって導かれる判定結果は、前述の従来の超音波診断や診断ガイドラインとは、全く別途の統計学的手法により幾何学模様の検出および判定を行っていることを忘れてはならない。教師学習用データ作成に失敗すると、これにより機械学習で得られたアルゴリズムの診断精度や病変検出能も低下する。
- 3) 1) および 2) などの理由により、超音波検査の教師学習用データは、病理や他の確定診断の得られた客観的データを元にして作成していくことが望ましい。しかし実際には、

病理学的結果の得られない良性疾患や稀な疾患も多いので、これらの検出・判定のための教師学習用データ作成は特に慎重になるべきである。

超音波診断装置は、ハードウェアの進歩により現在も小型化が進み携帯可能な画像診断装置として、ますます普及が見込まれている。しかしながら、装置を正しく操作してより高度な診断を遂行するための教育が現在課題となっている。より正確な診断や精度の高い病変検出結果を元に作成された機械学習アルゴリズムは、修練過程にある超音波検査の操作者の学習にも役立つことが期待される。また人工知能の超音波診断への応用は、前述した自動的ないしは半自動的に超音波画像を取得できる装置との組み合わせによって、より客観性のある検査に生まれ変わることに、さらなる装置の普及とそれによる低価格化が期待されている。さらに将来、人工知能の超音波診断への応用が進むと次に挙げる3つの可能性への発展が期待される。

- 1) 手術中に使用される超音波検査装置、いわゆる術中超音波は病変検出・診断精度の向上により、手術支援装置の一部として発展する可能性がある。
- 2) CT や MRI などと超音波検査画像の位置的重ね合わせ、いわゆる **fusion imaging** を用いた検査に人工知能による支援が加わる可能性が高い。
- 3) 2) と重複する部分もあるが、超音波検査同士、あるいは CT や MRI と組み合わせて経時的変化をより客観的に評価することにより、無駄な被曝や MRI など高額医用画像装置を用いた検査を省略できる可能性がある。

最後に、超音波検査装置はこれまで主としてハードウェアとして機械的に進歩をしてきた医療機器である。しかし今回のディープラーニングを中心とした人工知能の進歩は、ソフトウェアの進化にハードウェアとしてのコンピュータが追いついて実用化したといえる。超音波検査に従事する医療関係者や超音波装置開発者にとって、人工知能の超音波検査への応用に重要なのは、ハードウェアだけではなくソフトウェアをもっと重視する開発側の考え方といえる。

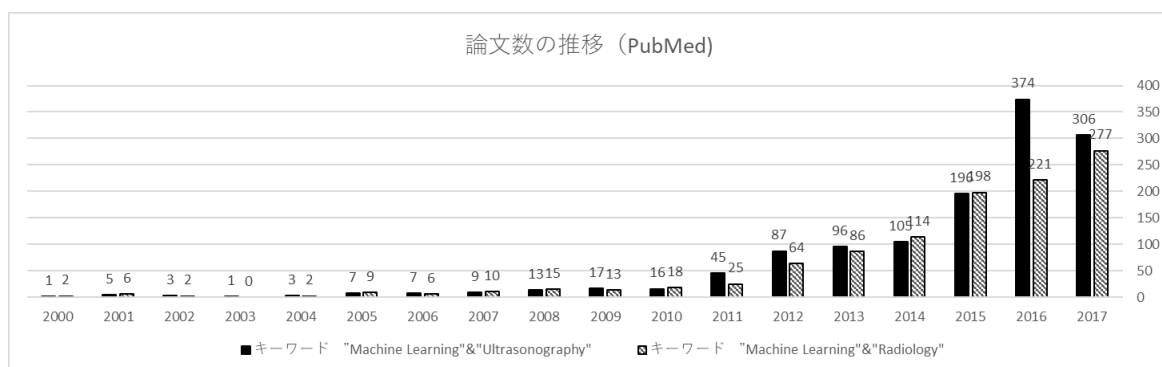


図1 超音波および放射線科領域の機械学習の論文数の推移

キーワード “Machine Learning” & “Ultrasonography” とキーワード “Machine Learning” & “Radiology” について PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/>) で各々検索した論文数の2000年から2017年までの年別推移を示した。

＜放射線診断分野における人工知能＞

国立研究開発法人 国立国際医療研究センター 国府台病院／日本医学放射線学会
待鳥 詔洋

エックス線は1895年にレントゲン（W.C.Roentgen）により発見され、発見後比較的早期から医療に応用されるようになった。我が国においては、国産の医療用X線装置第1号が1909年に千葉県市川市の国府台衛戍病院（現国立国際医療研究センター国府台病院）に設置され、放射線診断学が誕生した。その後、技術は年々高度化、精緻化し、各種造影剤の出現、フィルムや装置の改良など、多くのイノベーションが誕生した。放射線診断学は、フィルムに撮影されたいわゆるアナログ画像により発展するが、コンピュータが医療分野に応用されるようになると、画像のデジタル化が進み、画像はネットワークで配信されるようになった。また、新しいモダリティであるエックス線CT(Computed Tomography)、核磁気共鳴画像法(Magnetic Resonance Imaging : MRI)なども登場した。これらにより、情報量は飛躍的に増大し、空間分解能や時間分解能も著しく向上、形態だけでなく機能も画像化できる時代となった。

コンピュータによる診断支援を行う試みとしては、1980年代にはCAD (Computer-Aided DiagnosisまたはComputer-Aided Detection)が登場した。これは、放射線画像をコンピュータで定量的に解析し、その結果を「第2の意見」として利用する「診断の支援」をするシステムであった。単純エックス線写真から結節を特定するソフトウェアや乳房撮影（マンモグラフィ）の診断支援システムなどが開発され、臨床現場でも使われた。一方、2012年頃登場した「ディープラーニング」（深層学習：deep learning）は、診断支援を行うための機械学習技術の急速な進歩をもたらした。2012年に「ILSVRC (ImageNet Large Scale Visual Recognition Challenge)」でトロント大学のチームが勝利して以来、世界中のあらゆる分野で「ディープラーニング」を用いた技術革新がもたらされ、放射線診断学についてもさらなる技術革新が期待されている。

2016年、日本医学放射線学会は「Japan Safe Radiology」の概念を発表し、我が国の今後の放射線医学の発展のため、ICTを活用した構造改革を実施していくこととした。この中では人工知能（機械学習）による診断支援をさらに発展させていくことが盛り込まれた。また、2017年1月には日本放射線科専門医会がAI (Artificial Intelligence) 講演会を開催するなど、放射線診断に関連する医学系学会においても人工知能に関する期待が高まっている。

我が国においては、放射線科診断専門医がすべてのCT検査やMRI検査をレポートするためには、現在より少なくとも2.09倍の専門医数が必要との指摘があるなど¹⁾、多くのCT検査やMRI検査において放射線科診断専門医の関与が無いことが問題となっている。また一方で、社会保障関係費の増大が続く中、限られた医療資源を有効活用し、医療費の増大を抑制することが求められている。効率的でありながらも医療安全や医療技術をさらに向上させるため、人工知能（機械学習）技術の発展が期待されている。これらの難題を解決し国民の健康増進に寄与する画像診断支援システムが、数多く開発され、上市されることが望まれる。

参考文献

- 1) Nisie A., et al. Current radiologist workload and the shortages in Japan: how many full-time radiologists are required? Jpn Jpn J Radiol. 2015 May;33(5):266-72.

V. 参考資料

- 1 : 平成 23 年 12 月 7 日付け薬食機発 1207 第 1 号厚生労働省医薬食品局審査管理課医療機器審査管理室長通知「次世代医療機器評価指標の公表について」別添 3「コンピュータ診断支援装置に関する評価指標」
- 2 : 第 1 回会議 PMDA 講演資料「医療機器の承認審査について」
- 3 : 第 1 回会議 厚生労働省講演資料「人工知能を利用した医療機器の承認審査における課題 ～評価指標の位置づけと期待～」
- 4 : 第 1 回会議 東京大学 松尾 豊先生講演資料「人工知能は人間を超えるか –ディープラーニングの先にあるもの」
- 5 : 第 3 回会議 個人情報保護法委員会提供資料「個人情報保護法の概要」
- 6 : 第 3 回会議 内閣官房 健康・医療戦略室講演資料「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について」
- 7 : 第 18 回次世代医療機器・再生医療等製品評価指標検討会／医療機器開発ガイドライン評価検討会審査 WG 活動報告資料

薬食機発1207第1号
平成23年12月7日

各都道府県衛生主管部（局）長 殿

厚生労働省医薬食品局審査管理課
医療機器審査管理室長

次世代医療機器評価指標の公表について

厚生労働省では、医療ニーズが高く実用可能性のある次世代医療機器について、審査時に用いる技術評価指標等をあらかじめ作成し、公表することにより、製品開発の効率化及び承認審査の迅速化を図る目的で、検討分野を選定して評価指標を検討してきたところです。

今般、歯周組織治療用細胞シート、整形外科用カスタムメイド人工股関節及びコンピュータ診断支援装置の評価を行うに当たって必要と考えられる資料、評価のポイント等を評価指標としてとりまとめましたので、下記に留意の上、製造販売承認申請に当たって参考とするよう、貴管下関係業者に対しご周知いただきますよう御配慮願います。

なお、本通知の写しを独立行政法人医薬品医療機器総合機構理事長、日本医療機器産業連合会会長、米国医療機器・IVD工業会会長及び欧州ビジネス協会医療機器委員会委員長あて送付することを申し添えます。

記

1. 評価指標とは、承認申請資料の収集やその審査の迅速化等の観点から、製品の評価において着目すべき事項（評価項目）を示すものである。評価指標は、法的な基準という位置付けではなく、技術開発の著しい次世代医療機器を対象として現時点で考えられる評価項目を示したものであり、製品の特性に応じて、評価指標に示すもの以外の評価が必要である場合や評価指標に示す評価項目のうち適用しなくてもよい項目があり得ることに留意すること。
2. 個々の製品の承認申請に当たって必要な資料・データを収集する際は、評価指標に示す事項について予め検討するほか、可能な限り早期に独立行政法人医薬品医療機器総合機構の対面助言を活用することが望ましい。

コンピュータ診断支援装置に関する評価指標

1. はじめに

コンピュータ診断支援（Computer-Aided Diagnosis：以下 CAD）とは、X 線画像に代表される放射線画像をはじめとする医用画像に対して、コンピュータで定量的に解析された結果を「第 2 の意見」として利用する「医師による診断の支援」である。CAD の目的は、病変候補位置の情報をマーカで医師に示すことによって病変の検出を支援すると共に、がん病巣の悪性度等の病変候補の特徴に関する定量的なデータを医師に示して鑑別診断を支援するものである。

本評価指標は、コンピュータ診断支援装置の特に画像情報処理ソフトウェア部分について安全性と有効性を科学的根拠に基づいて適正かつ迅速に評価するために留意すべき事項を示すものである。

2. 本評価指標の対象

本評価指標は、医療用ソフトウェアのうち装置に付随するものであって、以下の CADe 又は CADx 機能を有する製品を対象とする。

開発するコンピュータ診断支援ソフトウェアが医療機器に該当するか判断し難い場合は、必要に応じ厚生労働省医薬食品局審査管理課医療機器審査管理室に相談すること。

(1) CADe (Computer-Aided Detection)

画像上で病巣が存在する候補位置をコンピュータが自動検出し、その位置をマーキングする機能を有するソフトウェアが組み込まれている装置であり、コンピュータにより医用画像データのみ又は医用画像データと検査データの両方を処理し、病変又は異常値の検出を支援するものである。

(2) CADx (Computer-Aided Diagnosis)

CADe 機能に加え、CADx は病変候補に関する良悪性鑑別や疾病の進行度等の定量的なデータを数値やグラフ等として出力する機能を有するソフトウェアが組み込まれている装置であり、診断結果の候補やリスク評価に関する情報等を提供して診断支援を行うものである。

3. 本評価指標の位置づけ

本評価指標は、技術開発の著しい医用画像工学・計算機科学等を対象とするものであることを勘案し、留意すべき事項を網羅的に示したものではなく、現時点で考えられる点について示している。よって、今後の更なる技術革新や知見の集積等を踏まえ改訂されるものであり、申請内容に関して拘束力を有するものではない。製品の評価にあたっては、個別の製品の特性を十分理解した上で、科学的な合理性をもって柔軟に対応することが必要である。なお、本評価指標の他、国内外のその他の関連ガイドラインを参考にすることも考慮すべきである。

4. 評価に対して留意すべき事項

(1) 基本的事項

- ① 開発の経緯、品目の仕様、国内外での使用状況、設計開発と装置の原理、目標とされる使用方法等（機能及び能力を含む）を明確に示す。
- ② 以下の事項を参考に、装置の設置、運用に当たっての留意事項等について評価する。

(ア) 設置

- a. プラットフォームの選択又は制限
- b. インストール及びアンインストールの条件と影響

(イ) 保守点検

- a. 周辺機器の稼動、必要資源の利用可能性
- b. ユーザ保守

(ウ) トレーニング計画の必要性とその内容

(エ) 実行環境の要求事項及び制限(複数ソフトウェアの共存)

- a. 使用リソース管理
- b. 実行環境の変更管理

(オ) 停電対策(無停電電源装置の要否等)

(カ) 電磁波等対策

(2) 非臨床試験

以下に示す各事項等を通して、ソフトウェアの安全性及び有効性の評価を適切に行うこと。

① 性能に関する評価

以下の各事項中、該当する項目について、それぞれ具体的なデータをもって明らかにすること。なお、使用データ属性については、モダリティ共通項目のほか、該当するモダリティ依存項目も評価すること。

(ア) 検出アルゴリズム（自動化の度合いを含む）

(イ) モダリティ共通の使用データ属性

a. 撮像機器

- ・製品（メーカー／機種／バージョン）

b. 撮影時条件

- ・造影剤：使用／不使用（使用の場合はその種類や使用量、撮影タイミング）
- ・撮影範囲：大きさと位置（詳細な条件の例：肝臓が全て撮像範囲内にある）
- ・空間分解能
 - 投影像：基準面でのピクセルの大きさ
 - 断層像：ピクセルの大きさ（スライス面内ピクセルの大きさとスライス厚及び間隔）

・時間分解能

撮像時間（写真のシャッタースピードに相当）、間隔（時系列画像のみ）

c. 画像出力

- ・デジタル画像信号深度：8ビット／16ビット
- ・デジタル画像化の際の信号ゲイン
- d. 保存状態
 - ・非可逆画像圧縮：有／無
- e. CAdE 処理
 - ・一回の検出処理に使用するデータ数：単一／複数（時系列、別条件、マルチモダリティ等）

（ウ）モダリティ特有の使用データ属性

- a. エックス線
 - ・撮影法：直接／間接
 - ・エックス線量（画像当たり単数／複数=DUAL ENERGY 等）
 - ・デジタル撮影／フィルムのスキャンデータ
 - ・その他必要な事項
- b. エックス線 CT
 - ・検出器：シングル／マルチ
 - ・スキャン：ノンヘリカル／ヘリカル
 - ・断層像再構成法（FBP/ART 等）及びその条件（再構成フィルタ等）
 - ・X 線量（画像当たり単数／複数=DUAL ENERGY 等）
 - ・その他必要な事項
- c. 超音波
 - ・周波数
 - ・走査形式（プローブ型：コンベックス／セクタ／リニア等）
 - ・モード（A/B/C/M/ドプラ等）
 - ・直接出力／間接出力（ビデオキャプチャ等）
 - ・その他必要な事項
- d. MRI
 - ・静磁場強度（1.5T／3.0T 等）
 - ・撮像種別（T1 強調／T2 強調／拡散強調等）
 - ・撮像法（Spin Echo 法／Echo Planar 法等）
 - ・パラメータ（TR/TE、信号収集回数等）
 - ・その他必要な事項
- e. その他のモダリティについても、撮像されたデータの画質に影響を及ぼす因子として、撮影時の条件を撮像装置固有のものも含めて詳細に記述すること。

（エ）検出性能

臨床画像データを用いた評価は、以下の事項等について行うこと。評価に使用する臨床画像データは、装置の目的や主要評価項目等を踏まえ、検出率や偽陽性・偽陰性率算出に必要なデータ数とし、複数施設のデータを含むものとする。臨床画像データを用いた方法では評価が困難な項目については、ファントム画像を用いて評価する。

- a. 絶対指標及び相対指標の種類、閾値及びその妥当性（ROC 曲線下面積、平均偽陽性数、判定能力等）

- b. 使用するデータセットの数、種類及びその妥当性
- c. 速度性能

② 安全性・品質に関する評価

次の設計管理に該当する事項を必要に応じて適用する。

(ア) 一般的要求事項

- a. 電氣的安全性
- b. 品質マネジメント
- c. リスクマネジメント

(イ) 仕様(稼働装置を含む)

- a. ニーズ (ユーザ、規格、規制、標準)
- b. インプット及びアウトプット (種類、方法、手段など)
- c. 個人情報保護
- d. セキュリティ
- e. 警報、警告及び操作者へのメッセージ
- f. OTS (Off-the-Shelf) ソフトウェアの特定
- g. ユーザビリティ
- h. 起動、中断及び終了
- i. インストール及びアンインストール

(ウ) 開発・設計計画

- a. 開発プロセス
- b. 成果物
- c. 開発プロセスを通して実行されたリスクマネジメントのトレーサビリティ
- d. リリース (複製、パッケージ等)

(エ) 検証計画

(オ) 構成管理計画

- a. OTS を含む管理対象アイテムの特定
- b. プロセスとツール

(カ) 文書化計画

- a. 文書開発、レビュー、承認及び改訂の手順と責任
- b. ユーザ向け文書 (操作及び保守マニュアル等)

参考：IEC 60601-1、ISO 13485、ISO 14971、IEC 62304、IEC/TR 80002-1、IEC 80001-1、ISO 12207、ISO 90003、IEC 61508-3 等

(3) 臨床試験(治験) (参考：平成 20 年 8 月 4 日付け薬食機発第 0804001 号)

① 治験の要否について

医療機器の臨床的な有効性及び安全性が性能試験等の非臨床試験成績又は既存の文献等のみによっては評価できない場合に臨床試験(治験)の実施が必要となり、臨床試験成績に関する資料の提出が求められる。また、その使用目的、性能、構造等が既存の医療機器と明らかに異なる場合については、原則として臨床試験の試験成績に関する資料の提出が必要である。

②有効性及び安全性の評価

治験を実施する場合においては、対象とする診断支援が実施できることを示すとともに、明確に定義された偽陽性や偽陰性を含む適切なエンドポイントを設定してその有効性と安全性を評価する。治験を実施する場合は、以下の事象に留意して実施計画を立案すべきである。

- (ア) コンピュータ診断支援ソフトウェアの特性
- (イ) コンピュータ診断支援ソフトウェアを適用する範囲
- (ウ) 使用する医師の習熟度の影響

③ 治験の症例数

臨床試験（治験）を実施する場合の症例数は、臨床試験の目的や主要評価項目等を踏まえ、科学的根拠に基づき、コンピュータ診断支援ソフトウェアの有効性や偽陽性・偽陰性率算出に必要な症例数とする。偽陽性・偽陰性率は別途行う臨床診断結果と比較して算出すること。

比較対照をおく場合にあっては、統計学的に症例数を設定する必要があることに留意すること。なお、対照として既存のコンピュータ診断支援ソフトウェアを使用する場合は、その妥当性について説明すること。

また、信頼できる海外データを承認申請の添付資料として使用できることがあるが、それのみで臨床評価を行うことができるかどうかについては十分に検討すること。



医療機器の承認審査について

独立行政法人
医薬品医療機器総合機構（PMDA）
医療機器審査第一部

本講演には、個人の立場の内容も含まれており、
PMDAを代表するものではありません。

アウトライン



1. 医療機器と製造販売承認
2. 承認審査の考え方
 - 1.1 臨床的位置づけ、有効性
 - 1.2 基本的な性能

1. 医療機器と製造販売承認

3

医薬品、医療機器等の 品質、有効性及び安全性の確保等に関する法律※

※ 「医薬品、医療機器等の品質、有効性及び安全性の確保等に関する法律」。
本スライドで「薬機法」とは、当該法を示します。

【目的】

- この法律は、医薬品、医薬部外品、化粧品、**医療機器**及び再生医療製品等の**品質、有効性及び安全性の確保並びにこれらの使用による保健衛生上の危害の発生及び拡大の防止のために必要な規制**を行うとともに、指定薬物の規制に関する措置を講ずるほか、**医療上特にその必要性が高い医薬品、医療機器及び再生医療製品等の研究開発の促進のために必要な措置を講じることにより、保健衛生の向上を図る**ことを目的とする。

（薬機法第1条）

4

【薬機法上での定義】

- 人若しくは動物の**疾病の診断**、治療若しくは予防に使用されること、又は人若しくは動物の身体の構造若しくは機能に影響を及ぼすことが目的とされている機械器具等（再生医療等製品を除く。）であって、**政令で定めるもの**

（薬機法第2条）



使用目的：	疾病の診断 、治療、予防 または 身体の構造、機能に影響を与える
条件：	政令で定められているもの

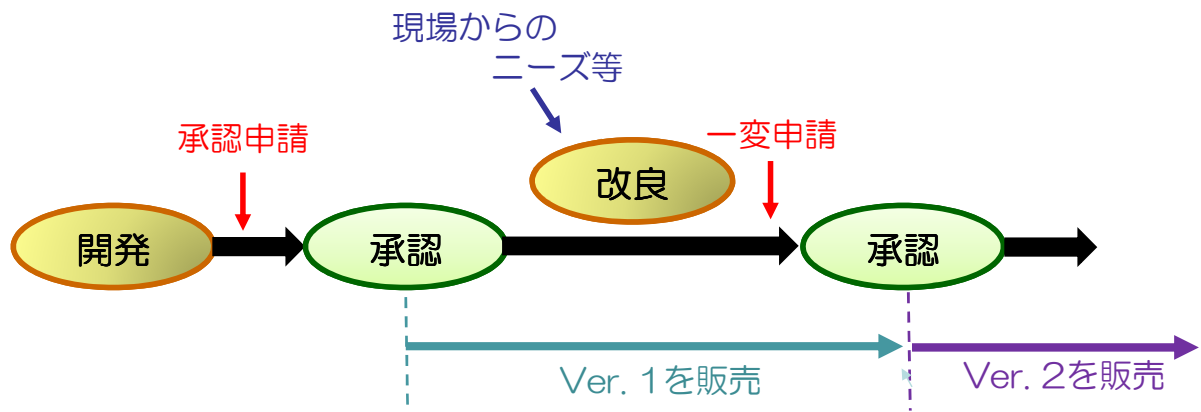
政令で定めるもの

- 機械器具
 - （84項目：例；医療用消毒器、呼吸補助器、内臓機能代用器、理学診療用器具、体温計、体液検査用器具、電気手術器、医療用はさみ、注射針、注射筒、歯科用ユニット、視力矯正法レンズ等）
- 医療用品
 - （6項目：例；エックス線フィルム、縫合糸、整形用品等）
- 歯科材料
 - （9項目：例；歯科用金属、歯冠材料等）
- 衛生用品
 - （4項目：例；月経処理用タンポン、コンドーム、避妊用具等）
- プログラム
- プログラムを記録した記録媒体
 - （6項目：例；疾病診断用プログラム、疾病診断用プログラムを記録した記録媒体等）
- 動物専用医療機器
 - （12項目：例；悪癖矯正用器具、受精卵移植用器具、製品蹄鉄等）

製造販売承認申請



- 医療機器の製造販売をしようとする者は、品目ごとに厚生労働大臣の承認を受けなければならない。(薬機法第23条の2の5 第1項)
- 承認事項の一部を変更しようとするとき(軽微変更を除く。)は、その変更について厚生労働大臣の承認を受けなければならない。(薬機法第23条の2の5 第11項)



※ 改良や変更があった際に、必ず申請が必要というわけではない。
(基本的には、有効性又は安全性に影響がある場合に申請が必要)

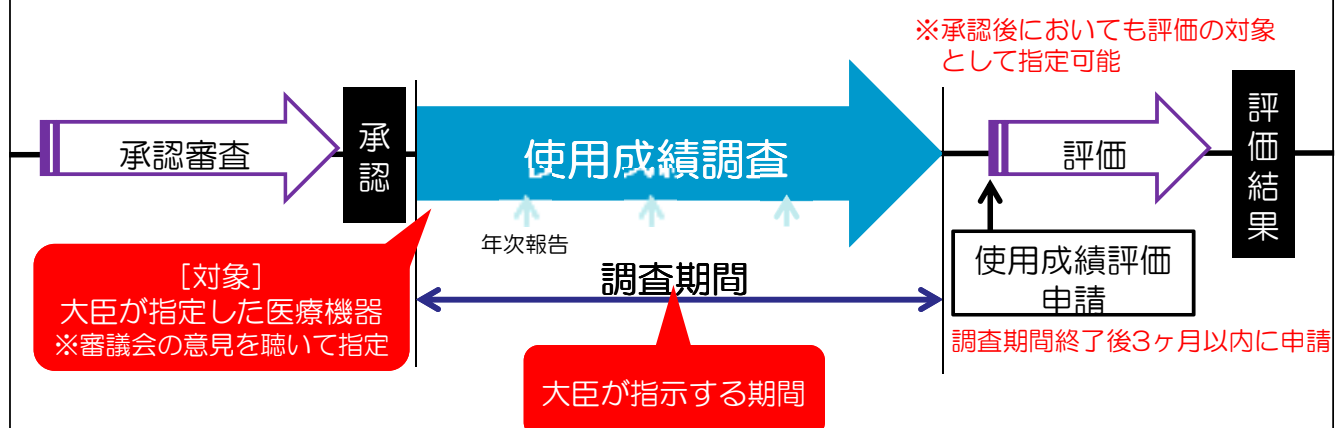
7

使用成績評価



薬事・食品衛生審議会の審議を経て承認を受けた者又は当該承認を受けている者は、その医療機器について、大臣が指示する期間を経過した日から起算して三月以内の期間内に申請して、使用成績に関する評価を受けなければならない。

(薬機法第23条の2の9 第1項)



8

診断機器を例に・・・

2. 承認審査の考え方について

9

承認拒否事由

- ・ 医療機器の種類に応じた製造販売業の許可を受けていないとき
- ・ 当該医療機器を製造する製造所が製造業の登録を受けていないとき

- ・ **申請品目の審査の結果、次のいずれかに該当するとき**

- ① 効果又は性能を有すると認められないとき
- ② 効果又は性能に比して著しく有害な作用を有することにより、使用価値がないと認められるとき
- ③ 医療機器として不適当と厚生労働大臣が定めたものに該当するとき。

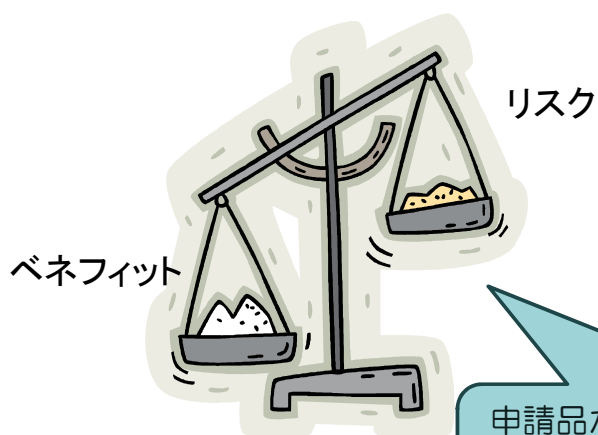
- ・ 医療機器の製造所における製造管理・品質管理の方法が、QMS基準に適合していると認められないとき

(薬機法第23条の2の5)

承認審査のポイント： 有効性と安全性のリスクベネフィットバランス



審査では、申請資料に基づき、
有効性・安全性のリスクベネフィットバランスを評価し、
使用目的又は効果、使用方法、
使用上の注意の妥当性が判断される。



- 臨床現場において、申請品はどのような（どの程度の）効果・結果が期待されているか？
 - 申請品はその性能を有しているか？
 - その性能を発揮するためにどのようなリスクが介在するか。
 - そのリスクは許容できるのか？
- etc……

申請品がこのようになっていることを、
申請者が根拠（試験結果等）と共に
申請資料で説明する。

11



2.1 臨床的位置づけ、有効性

12



申請者



この装置は、
がん診断に使用します！

従来のがん診断の中で、どういうフェーズで使うのですか？
新しいフェーズを生み出すものですか？
既存の診断方法に置き換わるものですか？

- どういう条件でどれくらいの診断能があれば、有効なものと言えるだろうか・・・
- どれくらいの診断エラーなら許容できるだろうか・・・



PMDA

13

【CASE1: 既存の診断スキームに置き換えようとする】



スクリーニング
(通常年1回)

確定診断

治療
(副作用重篤)

- たくさんの陰性のなかから、陽性を見つけることができるか？
- 偽陰性と偽陽性、どちらがリスクとしては高いだろうか。

- 絞り込まれた群の中から、正しく陽性、陰性を判断できるか？
- 偽陰性と偽陽性は、どれくらい許容できるか。

14

臨床的位置づけと評価



【CASE2: 既存の診断スキームに新たなスキームを加える】



スクリーニング
(通常年1回)

本品による
診断

確定診断

治療
(副作用重篤)

- 確定診断にいかずに、この診断をやるメリットは何だろうか？
(確定診断よりも低侵襲、早い etc...))
- メリットに対して、どの程度の診断性能（感度、特異度、偽陽性率、偽陰性率など）があれば、臨床的に価値があるといえるだろうか？
- このフェーズで誤診をすることのリスクはどんなものだろうか？

15

臨床的位置づけと評価



【CASE3: 新たな診断を作る？】



本品による
診断

治療は研究中
(まだない)

この診断をする意義は何だろうか？



16

審査の考え方①

意図した臨床上の意義を達成できているか？



臨床上の位置付けによって、示すべき有効性、安全性のポイントが変わる。

例) 診断支援プログラム

スクリーニング
(通常年1回)

確定診断

治療
(副作用重篤)

- 多くの陰性を含む集団から、陽性を見つけられることが求められる。
- 偽陰性と偽陽性のリスク、プログラムの位置付け等により異なる。
(例えば、確定診断に行くべき機会を失うならば、偽陰性のリスクは高い。)

- 陽性疑いの群の中から、正しく陽性、陰性の判断ができることが求められる。
- 偽陰性のリスク：治療機会を失う
- 偽陽性のリスク：不要な治療介入

“感度及び特異度が100%を達成”は現実的ではない。
偽陽/陰性となりやすい条件等も理解し、情報提供すべき内容と合わせて審査する。

17

審査の考え方②

性能向上＝有効性向上とは限らない



従来のモダリティより、画像がきれいになりました！
⇒ だから、既存のものより高い有効性が発揮できます！！



従来モダリティ



新モダリティ1



新モダリティ2

確かに、画像は
きれいになったが・・・

診断結果に差が出るくらいの改良になった？
患者にとってどれくらいメリットになった？

臨床上のアウトカムに対する影響で評価

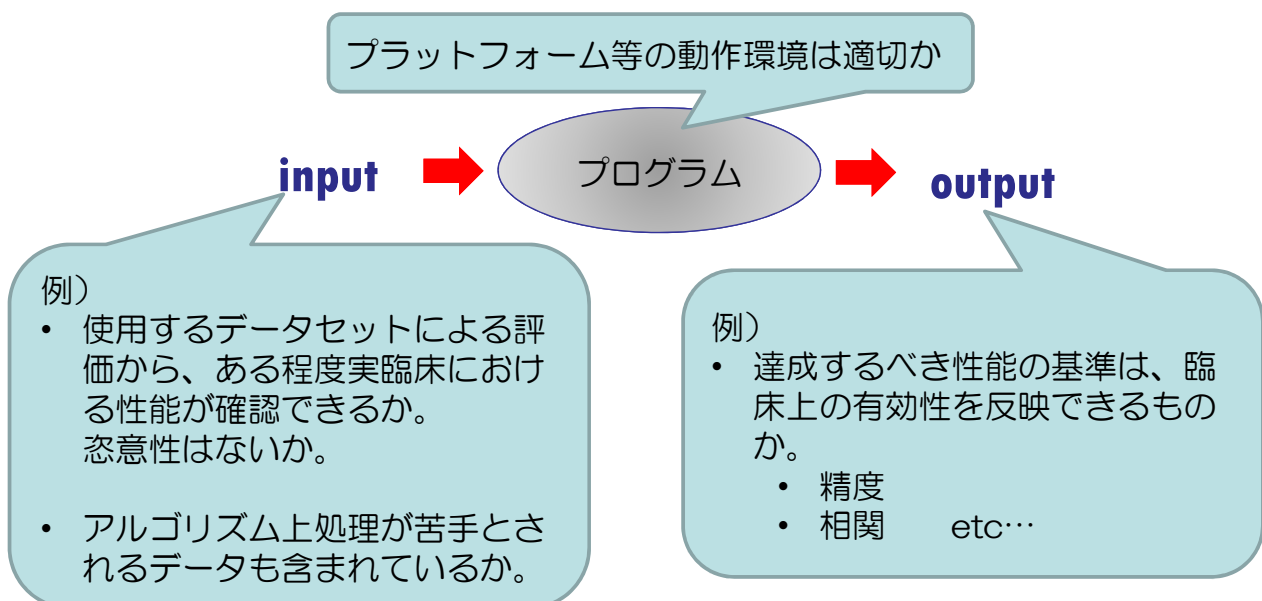


18

2.2 基本的な性能

19

審査の考え方③ 基本的な性能の評価は適切であるか？



臨床現場に出した際に実力が発揮できる程度の性能があることを評価してもらう。

20

医療機器プログラムの 承認申請に関するガイダンス



医療機器プログラムの承認審査の経験から、
審査に際して必要な評価の論点をまとめた文書
(H28.3.31)

4. 審査上の論点

(1) コンセプト、機能について

- ① 臨床的意義
- ② 機能の把握と特定
- ③ アルゴリズムの明確化
- ④ 使用環境要件
- ⑤ 併用される医療機器等の特定

臨床的位置づけ、有効性

基本的な性能

(2) 評価について

- ① アルゴリズムの妥当性、臨床的意義を踏まえた評価
- ② 検体のバージョン管理
- ③ 比較対照の妥当性
- ④ 入出力の妥当性
- ⑤ 精度の評価
- ⑥ 実試験との相関の評価、シミュレーションの妥当性

21



ご清聴ありがとうございました



人工知能分野審査WG 平成29年度第1回会議

人工知能を利用した医療機器の承認 審査における課題 ～WG評価指標の位置づけと期待～

厚生労働省 医薬・生活衛生局
医療機器審査管理課

1

AIの活用への期待の高まり

- 未来投資会議(官邸)
- 人工知能と人間社会に関する懇談会(内閣府)
- AIネットワーク社会推進会議(総務省)
- 人工知能技術戦略会議(文科省)
- 保健医療分野におけるAI活用推進懇談会(厚労省)
- 新産業構造ビジョン(経産省)

2

AI技術を用いた医療機器への対応について

AI技術を用いた製品のうち、その使用目的や提供形態等から医療機器に該当するものは、医薬品医療機器法に基づき安全性、有効性の確保が行われる。

医療機器規制

研究開発、臨床試験

承認審査

製造・品質管理

流通、市販後安全対策

安全性・有効性等の評価

これまでの対応

- ロボット技術、ICT等の技術革新を踏まえ、PMDAの医療機器に係る相談・審査体制を再編（平成27年10月）
- 診断の支援等を行う医療機器プログラムの審査上の論点（臨床意義を踏まえた評価、実試験との相関等）をまとめたガイドランスを公表（平成28年3月）
- 画像診断機器メーカー等とAIの活用に関する意見交換会を実施。引き続き実施予定。（平成28年12月～）

課題

1. 評価の手法

- 継続的な性能の変化などのAIの特性により即した形で医療機器としての評価を行う必要がある。
→ まずは、早期の実用化が期待される画像を用いた診断分野に着目し、**AI技術を活用する画像を用いた診断機器の評価指標等**について検討を行う予定。

2. 評価の体制

- 円滑な実用化には、開発相談、審査に加え、市販後の評価・フォローアップが重要。継続的な性能の変化などのAI技術の特性・可能性を理解して、継続的に適切な評価や医療現場への情報提供等を行うことが求められている。
→ 開発の進展に応じて、**医療機器の市販前・市販後の評価に係る体制整備**を検討すべきではない

医療機器プログラムとは

医療機器プログラム

汎用PC等にインストールすることで、医療機器としての性能を発揮するプログラム※

プログラムを使用している医療機器

画像診断装置ワークステーション



X線CT装置、MRI、PET-CT装置等で撮影された画像データの処理、保存、表示等を行う

<使用例>



CT検査のデータ

当該製品のプログラムによる3次元画像処理



骨の3D画像

旧薬事法



ソフト部分
(プログラム)



ハード部分

ソフト部分のみでは薬事法の規制対象とならず、ハード部分に組み込んだ形で規制していた

法改正

単体プログラム



ソフト部分
(プログラム)

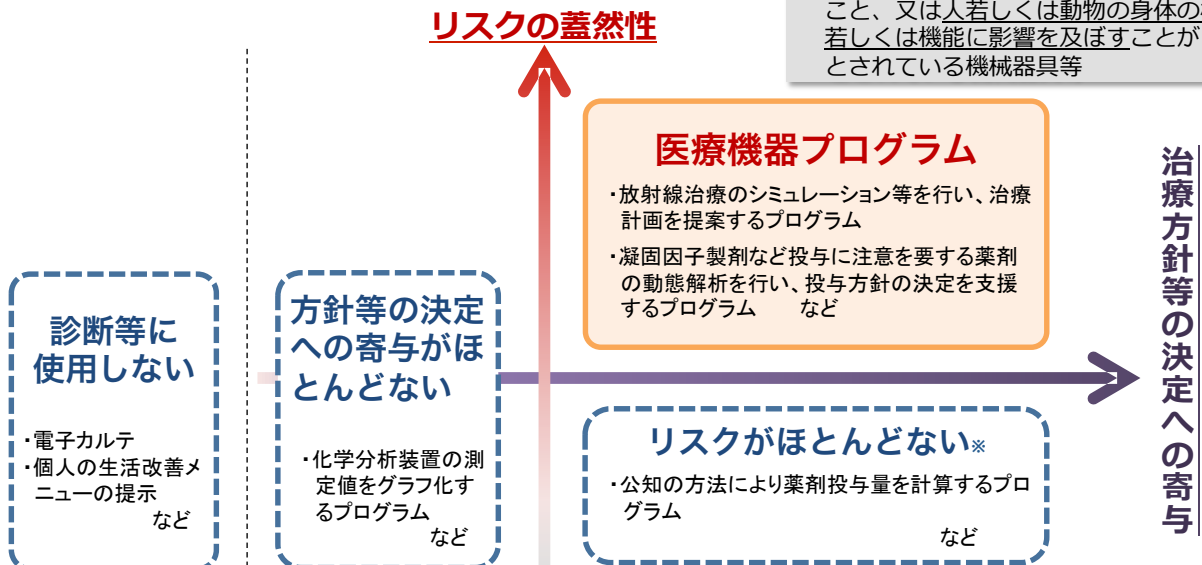
**プログラム単体で
薬事法の規制対象とした**

※ 欧米では、従前から医療機器として位置付けられている。

医療機器プログラムの該当性の考え方

- **医療機器の定義※に合致するプログラム**が該当。ただし、機能の障害等が生じた場合でも**人の生命、健康に影響を与えるおそれがないものは除外**。
- 個々のプログラムが医療機器に当たるかどうかは、①**治療方針等の決定への寄与の大きさ**、②**不具合が生じたときのリスク**等を勘案して判断。

※**医療機器の定義** 人若しくは動物の疾病の診断、治療若しくは予防に使用されること、又は人若しくは動物の身体の構造若しくは機能に影響を及ぼすことが目的とされている機械器具等



※ 仮に機能の障害があった場合でもリスクが低いもの。
例えば、使用する医師が容易に間違いを認知できるものなど。

5

医療機器プログラムの提供とサービス行為の違い

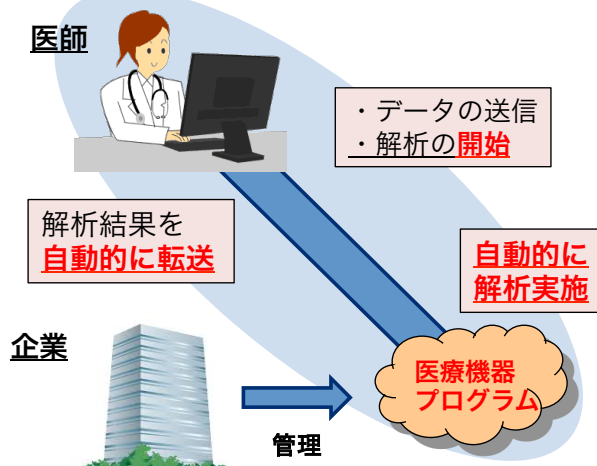
プログラム医療機器の提供方法には、**ネットワーク上に存在するプログラム医療機器の使用権（アクセス権）を提供する形式**も存在する。

※ **プログラム医療機器の使用権の提供は、医療機器の製造販売とみなす。**

（以下通知抜粋） 電気通信回線を通じた医療機器プログラムの提供には、ダウンロード販売に加え、医療機器プログラムの所有権は移転せずに使用権を認める形態が含まれること。

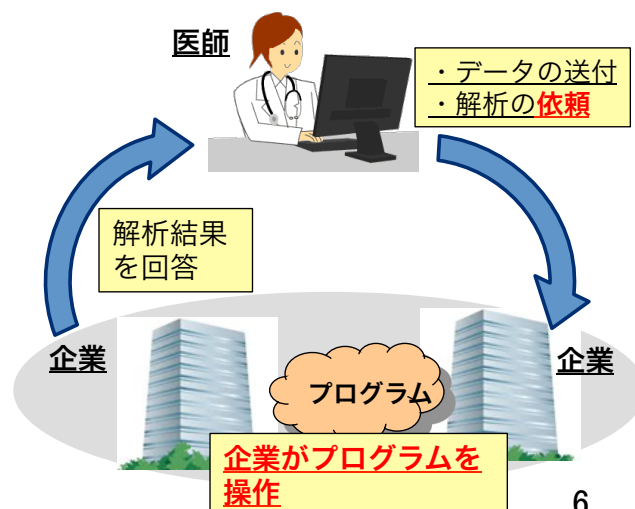
医療機器プログラムの提供 （使用権の提供）

→ （原則）**医師が直接プログラムにアクセス**



解析サービスの提供 （受託解析）

→ 医師の依頼により**企業がプログラムを操作**



6

A I による診療支援と医師の判断との関係性の整理（案）

保健医療分野におけるA I 活用推進懇談会資料より

構成員からのご意見

- 最終意思決定（診断確定、治療方針確定）は医師が行うことを明確化しておく必要はないか。
- A I の推測を基にした診断に誤りがあった場合に責任を負う者を明確化する必要はないか。

（第1回「保健医療分野におけるA I 活用推進懇談会」でのご意見より）

現状

- A I が単独で診断、治療方針の決定を行っているわけではなく、A I の推測結果を、他のデータベースの情報と統合して医師が解析し、治療方針の決定を行っている。
- A I の推測結果が誤っていた場合、研究者が間違いを指摘し改良を行っている。
- 画像診断に関しても、A I が数学的な特徴抽出を行うことで推測した結果を、医師が解釈し、患者に説明することが必要。

方向性

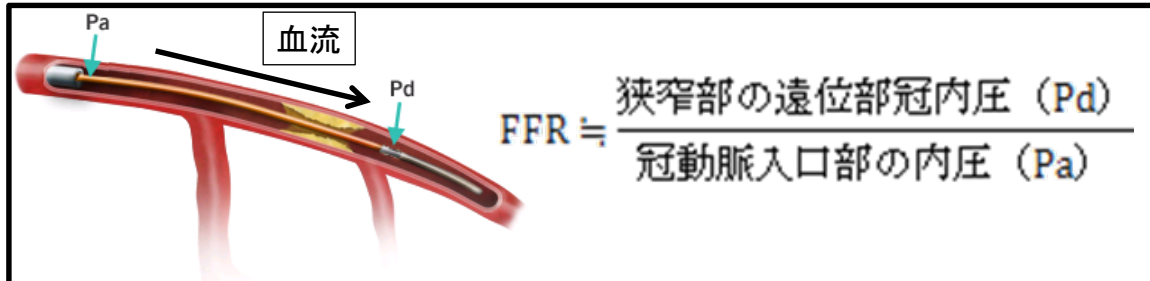
- 現状ではA I が単独で診断確定や治療方針の決定を行うことはできないため、これらの**最終的な意思決定は医師が行う**こととしてはどうか。
- 現状ではA I の推測結果には誤りがありうるが、この場合、A I を活用し最終的な意思決定を行う**医師が責任を負うべき**ではないか。その前提として、**医師に対してA I についての適切な教育**を行い、安全性を確保していくことが必要ではないか。
- A I を活用したより良い診療支援の確立のためには、**保健医療分野におけるA I 開発への医師の関与**が必要ではないか。

ハートフローFFR_{CT}

- ・ 類別：疾病診断用プログラム
- ・ 一般的名称：循環動態解析プログラム
- ・ 販売名：ハートフローFFR_{CT}
- ・ 申請者：ハートフロー・ジャパン合同会社
- ・ 申請年月日：平成27年2月19日
- ・ 承認年月日：平成28年11月14日
- ・ 概要：本品は冠動脈疾患の虚血（心筋への血流量が不足している状態）を評価するための支援プログラムである。

FFR

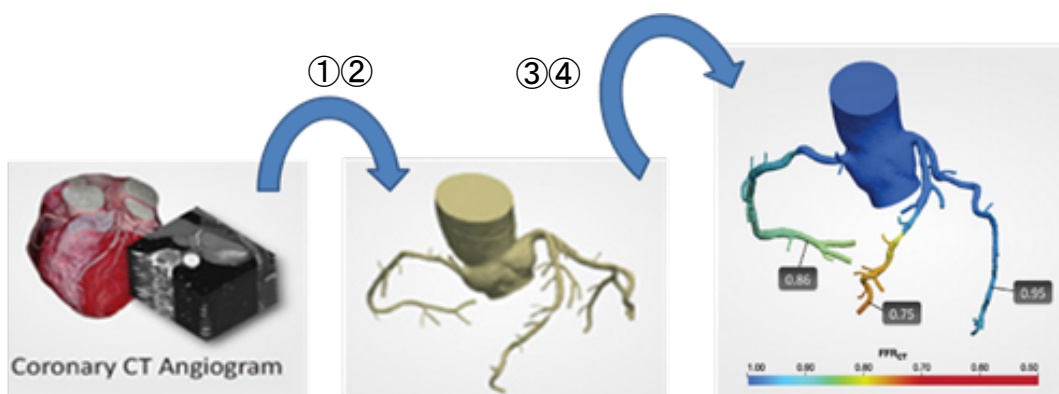
- 虚血を確認する検査方法の1つとして、冠血流予備量比 (FFR:Fractional Flow Reserve)が挙げられる。
- FFRは、圧測定ワイヤを**経皮的に冠動脈へ直接挿入**し狭窄病変の前後の圧を測定することで、虚血の評価が可能である。→**侵襲的！**



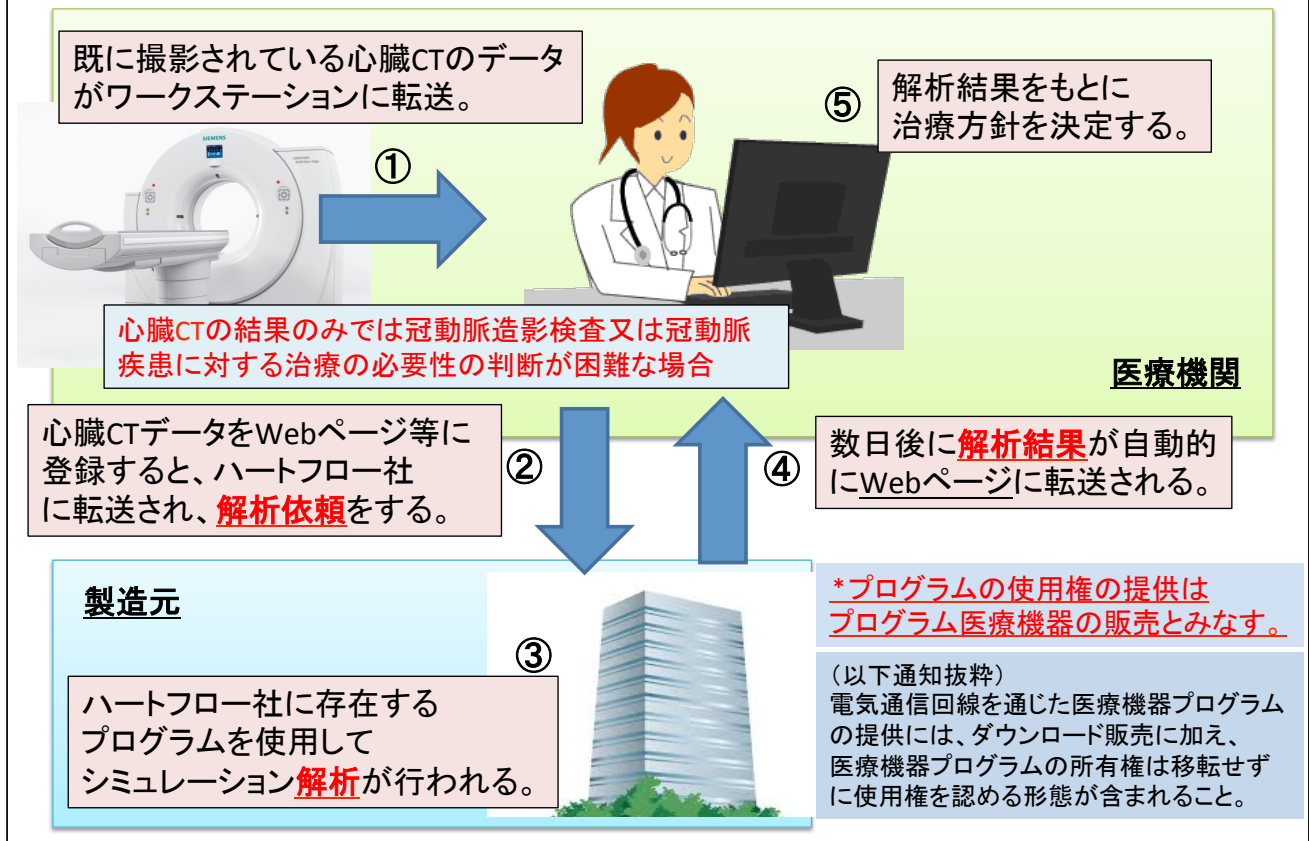
- ハートフローは、冠動脈疾患の疑いのある患者の診断に際し、(既に撮影された)**心臓CTデータから 非侵襲的に解析した仮想FFRをFFR_{CT}として提供**することにより、機能性虚血の評価を支援するためのプログラム。

FFR_{CT}のシミュレーション

- ①心臓CTデータから、当該患者の冠動脈の構造を再現する。
- ②冠動脈及び脈管構造より生理学的数値*を算出する。
 - *血液流入量: 冠動脈への血液流入量は冠動脈の総体積または心筋体積に比例することを用いて血液流入量を算出する。
 - *血液抵抗値: 冠動脈入口部の圧を血液流入量で除することにより算出する。
 - *流体/パラメーター: 血液粘度 ($5.34902 \times 10^{-3} \text{g}/(\text{mm} \cdot \text{s})$)、血液密度 ($0.00106 \text{g}/\text{mm}^3$)
- ③冠動脈内に**仮想の血流を流す事で、冠動脈内圧を算出***する。
 - *ポアズイユの法則やベルヌーイの定理に従う減次モデルや三次元ナビエ・ストークス方程式を適用する。
- ④狭窄病変前後の圧の比によりFFR_{CT}値を算出する。



流通形態



ハートフローFFR_{CT}

承認条件

本品を使用する医療機関から製造元に転送されたCT データ等がFFRCT 値の算出又は本品の必要な改良改善、及び次世代品の開発以外の目的に使用されないよう必要な手続き・適切な管理を行うとともに、不法なアクセスを防止するため最新のセキュリティ対策を講じること。

指示事項

- ✓ 臨床試験成績に基づき、本品の感度及び特異度について、添付文書及びトレーニングプログラムで注意喚起を行う。
- ✓ 本品により算出されるFFRCT値は実測FFR 値と必ずしも一致するものではないことから、本品の解析結果のみにより診断しないことを添付文書等において注意喚起する。
- ✓ FFRCT の診断性能は、偽陽性が多めに検出される傾向が示されたため偽陽性及び偽陰性に関する情報について添付文書及びトレーニングプログラムで注意喚起を行う。
- ✓ 本品の解析は使用される CT 画像の質に影響を受けるため、解析可能なCT 画像の要件やその撮像条件等について、添付文書及びトレーニングプログラムにより周知する

適正使用指針

関連学会に依頼し、適正使用に関するガイドラインを作成

- ✓ ハートフローの適応となる患者の基準
- ✓ 使用する医師の条件
- ✓ 使用する施設の条件
- ✓ 個人情報保護

等

これまでに示された プログラム医療機器の申請に 関するガイダンス

13

プログラム医療機器の申請に関するガイダンス

平成28年3月31日医療機器再生医療等製品担当参事官室事務連絡

審査上の論点

- － プログラムのコンセプト、機能の実体に関する事項
 - ・ 臨床的意義の明確化
 - ・ 機能の把握と特定（入力/出力）
 - ・ 計算アルゴリズムの明確化
 - ・ プラットフォーム、使用環境の要件の特定
 - ・ 併用機器/医薬品等の条件の特定
- － 評価に関する事項
 - ・ アルゴリズムの妥当性、臨床的意義を踏まえた評価
 - ・ 試験検体のバージョン管理
 - ・ 比較の対象の妥当性
 - ・ 入出力データの妥当性
 - ・ 精度の評価
 - ・ 解析結果と実試験の相関関係の評価/シミュレーションの精度の評価
- － 承認後の変更に関する事項

14

これまでに示された 診断支援装置についての評価指標

15

コンピュータ診断支援装置評価指標

平成22年度
次世代医療機器評価指標作成事業

コンピュータ診断支援装置
審査WG報告書

平成23年3月

審査WG座長 遠藤 啓吾
群馬大学大学院医学系研究科
放射線診断核医学

16

コンピュータ診断支援装置評価指標

評価に対して留意すべき事項

- － 基本的事項
 - ・ 開発の経緯、品目の使用、設計開発と装置の原理、使用方法（機能及び能力）等の明確化
 - ・ 装置の設置、運用に当たっての留意事項（プラットフォーム、保守点検、トレーニング計画、実行環境の要求事項等）
- － 非臨床試験
 - ・ 性能に関する評価（検出アルゴリズム、モダリティ共通/特有の使用データ属性、検出性能）
 - ・ 安全性・品質に関する評価（一般的要求事項、仕様、設計計画、検証計画、文書化計画）
- － 臨床試験
 - ・ 治験の要否についての検討
 - ・ 有効性及び安全性の評価（CADソフトウェアの特性、適用範囲、使用する医師の習熟度の影響等に留意）
 - ・ 治験の症例数の検討（対象の妥当性）

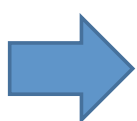
17

AIを用いた画像診断機器の評価指標

・ ターゲット

「AIを用いた医療機器」は幅広く可能性がある

- － AIを用いて開発されたアルゴリズムを組み込んだ医療機器
- － 診療記録や検査値等の患者情報を入力すると関連する情報を提示し診療を支援する医療機器
- － ゲノム等患者のデータを解析し関連情報を提示し診断を支援する医療機器
- － 患者のデータを入れると診断名を表示してくれる医療機器...

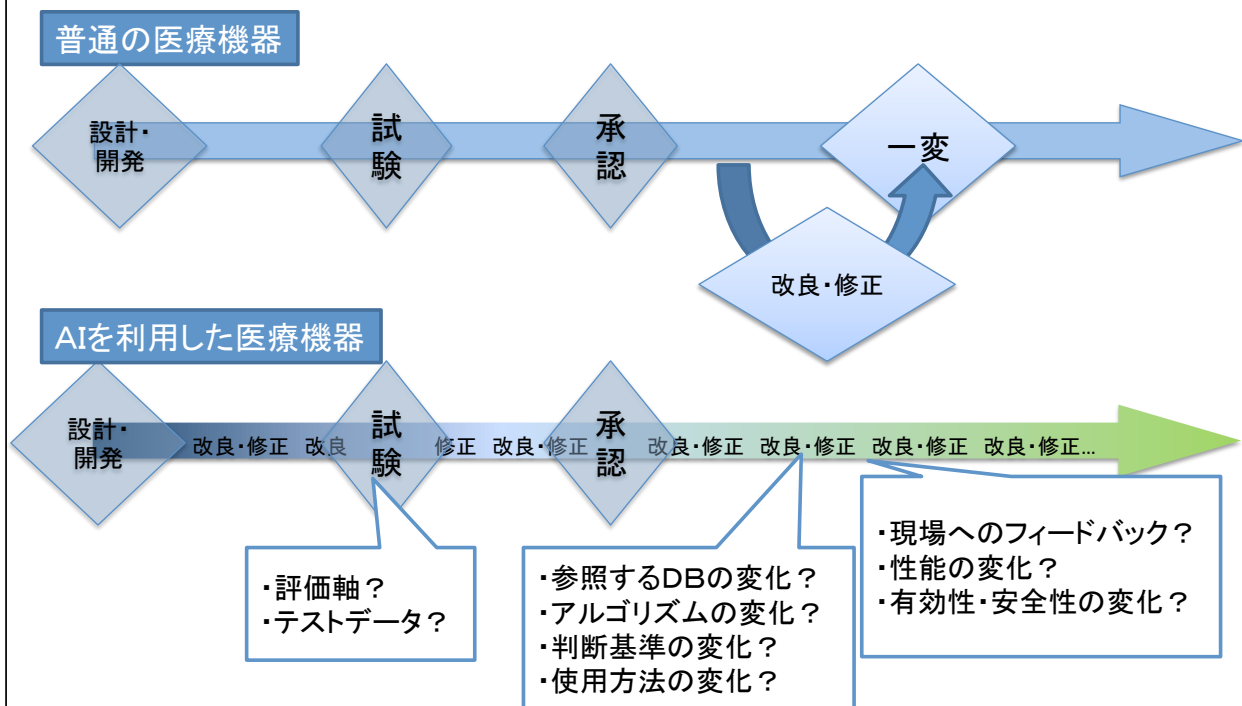


今回は、画像を解析し、医師の見落とし防止や一次スクリーニング等の診断支援を行うAIを想定して検討してはどうか

18

AI特有の懸念

これまでにガイドが示された医療機器プログラムや診断支援装置との違い



19

AI特有の検討事項の候補

- ・ 性能の評価、性能変化の検証
 - － 評価軸、検証データ、対照、テストデータ
- ・ 性能変化の管理
 - － 日常診療下での検証、施設ごとの変化
- ・ 問題発生時の対応
- ・ 性能の変化による使用目的の承認からの逸脱
- ・ 使用者への注意喚起、トレーニング

...などが考えられるのではないかと

20

人工知能は人間を超えるか

- ディープラーニングの先にあるもの

東京大学 松尾 豊

1

東京大学 松尾研究室について



松尾 豊

1997年 東京大学工学部電子情報工学科卒業
 2002年 同大学院博士課程修了。博士(工学)。産業技術総合研究所 研究員
 2005年 スタンフォード大学客員研究員
 2007年～ 東京大学大学院工学系研究科 技術経営戦略学専攻 准教授
 2014年～ 東京大学 グローバル消費インテリジェンス寄付講座 共同代表・特任准教授
 2015年～ 産総研AIセンター 企画チーム長

- ◆人工知能、ディープラーニング、Webマイニングを専門とする。
- ◆論文数と被引用数に基づき科学者の科学的貢献度を示すh-Index=31(ウェブ・人工知能分野最高水準)であり、2013年より国際WWW会議Web Mining部門のチェアを務める。
- ◆世界人工知能国際会議 プログラム委員。2012年より、人工知能学会 理事・編集委員長(それまでの慣例を大幅に更新し最年少で編集委員長就任)、2014年から倫理委員長。
- ◆人工知能学会論文賞(2002年)、情報処理学会長尾真記念特別賞(2007年)、ドコモモバイルサイエンス賞(2013年)、文部科学省 科学技術への顕著な貢献2015、大川出版賞(2015年)、ビジネス本大賞審査員賞(2016年)等受賞。
- ◆経済産業省 産業構造審議会 新産業構造部会 委員、IoT推進コンソーシアム 運営委員、厚生労働省「働き方の未来 2035」懇談会メンバー、内閣府「人工知能と人間社会に関する懇談会」構成員、金融庁「フィンテック・ベンチャーに関する有識者会議」委員、総務省「ICTインテリジェント化影響評価検討会議」委員等。
- ◆近著に「人工知能は人間を超えるか?—ディープラーニングの先にあるもの」(角川 2015)。

<研究室の実績>

- ◆博士学生17人、修士・学部生10人が所属し、人工知能の基礎研究、ソーシャルメディアの分析、データ分析及びその実社会へのアプリケーションを多方面にわたって行っている。
- ◆これまでに、トヨタ、リクルート、マイクロソフト、CCC、経営共創基盤、ミクシィなどさまざまな企業と共同研究の実績がある。官公庁からも、経産省(アジアトレンドマップ等)、文科省(ビッグデータ活用)など相談多数。
- ◆卒業生の主な進路は、Google、DeNA、楽天、サイバーエージェント、光栄、ゴールドマンサックス、BCG、三井物産、電通など。起業した学生も多数。GunosyやREADYFORなどのサービスを構築、運用している。

Googleの人工知能(アルファ碁)が囲碁でプロ棋士を破る (2016年3月)

カルチャー

文字サイズ 小 中 大

囲碁人工知能、「深層学習」で最強棋士に3連勝

2016年03月13日 07時13分 Tweet



人工知能アルファ碁と対戦する李九段(右) = A P

【ソウル=川村律文】米グーグル傘下の英グーグル・ディープマインド社が開発した囲碁の人工知能「アルファ碁」と世界のトップ棋士である韓国棋院のイ・セドル 李世九段(33)の第3局が12日、ソウル市内のホテルで行われ、アルファ碁が勝利した。(●は、石の下に乙)

5局まで行われる対戦で、アルファ碁は3連勝を果たし、最強棋士との対決を制した。チェス、将棋に続き、頭脳ゲー

ネット上の謎棋士「マスター」、正体はアルファ碁進化版

大出公二、深松真司 2017年1月5日07時16分

シェア 273 ツイート list 32 ブックマーク メール 印刷



昨年、通算成績4勝1敗でアルファ碁に敗れた韓国の李世九段(右)は「石」の下に「乙」、イセドル九段(左)は「石」の下に「乙」、イセドル九段(右)は「石」の下に「乙」、イセドル九段(左)は「石」の下に「乙」。



年末年始、ネット上の中国の囲碁サイトにハンドルネーム「Master(マスター)」なる棋士が参戦し、非公式ながら世界のトップ棋士とみられる対戦相手に対し、今月4日までの1週間で60勝無敗という驚異の戦績を上げた。その正体は、昨年韓国の世界トップ棋士を破った囲碁A I(人工知能)の「アルファ碁」の進化版だったことが明らかにされた。

囲碁サイトには少なからぬプロ棋士がハンドルネームを用い、短時間の非公式対局で腕試しをしているが、トップ棋士のハンドルネームは半ば公然と囲碁ファンに知られている。

英語で「名人」「達人」などの意味を持つMasterは、12月29日に囲碁サイト「東洋囲碁」、元日には「野狐囲碁」に登場。現在世界ナンバー1と評される中国の柯潔九段や韓国の朴廷桓(パクジョンファン)九段、日本の井山裕太六冠ら世界のトップ棋士とみられるハンドルネームの対戦相手をハンディなしの互先(たがいせん)で次々に破った。

読売新聞2016/3/13、朝鮮日報2016/3/11

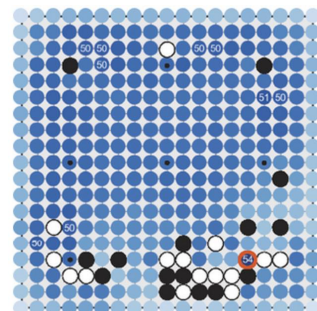
関係者の衝撃

思考ゲームの歴史

- 1980年: オセロプログラムMoorが、世界チャンピオン井上博との六番勝負で1勝を挙げた
- 1997年: チェスプログラムDeep Blue (IBM) が、世界チャンピオン(ガルリ・カスパロフ)に勝つ
- 2012年: 将棋プログラム ボンクラーズが、故・米長永世棋聖に勝つ
 - ・ その後、2015年まで、プロ棋士と対局し、9勝5敗1分
- 残るは囲碁のみ
 - ・ 「将棋の10年遅れ」でまだ10年は大丈夫。解の空間が広く、難しいため。
 - ・ 2015年当時、まだアマチュアレベル。日本は研究でリード

Google(DeepMind)のアルファ碁が、トッププロ棋士に勝利

- 2015年1月にNature誌に掲載
- Googleが買収したDeepMindが参入し、あっという間に(1年程度で)抜かれてしまった
- 3月にトッププロの韓国のイ・セドルと対局。4勝1敗でアルファ碁の勝利。
- 12月には、囲碁サイトに現れ、60連勝
- 鍵となるのは、ディープラーニングを活用し、局面の「認識」技術を使ったこと



人工知能をめぐる動向

第1次AIブーム(1956～1960年代):探索・推論の時代

- ダートマスワークショップ(1956)
 - ・ 人工知能(Artificial Intelligence)という言葉が決まる
 - ・ 世界最初のコンピュータENIAC(1946)のわずか10年後
- 数学の定理証明、チェスを指す人工知能等

考えるのが早い人工知能

...冬の時代

第2次AIブーム(1980年代):知識の時代

- エキスパートシステム
- 医療診断、有機化合物の特定、...
- 第5世代コンピュータプロジェクト:通産省が570億円

ものしりな人工知能

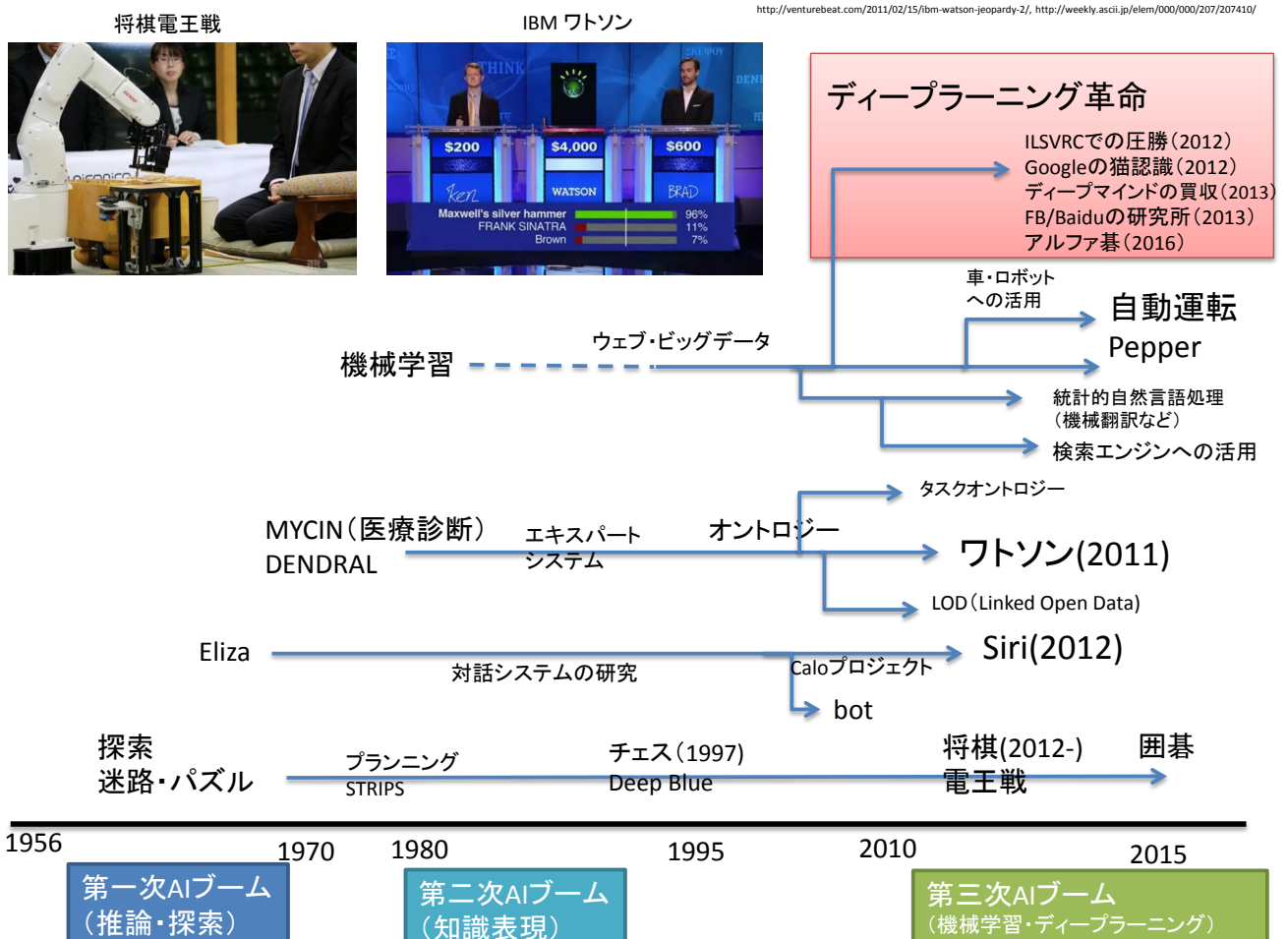
...冬の時代

第3次AIブーム(2013年～):機械学習・ディープラーニングの時代

- ウェブとビッグデータの発展
- 計算機能力の向上

データから学習する人工知能

5



ディープラーニング革命

認識

「画像認識」ができる
(コンピュータができて以来、初めて！)



運動の習熟

ロボット・機械に
熟練した動きができる



言葉の意味理解

文の「意味」が分かる
(文と映像の相互変換ができる)

7

認識の難しさ



ネコ



イヌ



オオカミ

- これをコンピュータで見分けたい
 - 目が丸い→ネコ
 - 目が細長い、耳が垂れている → イヌ
 - 目が細長い、耳がとがっている → オオカミ



イヌ



結局、「耳が垂れている」「目が細長い」などの「特徴量」を人間が考えている限り無理。どんなに頑張っても、必ず例外がある。人間はなぜかうまくできる。

8

これまでの人工知能の壁＝特徴抽出の壁

- 難しい問題1: 機械学習における特徴量の設計 (Feature engineering)
 - － 機械学習において、変数(特徴量)の設計が難しかった。
 - － 人間が対象をよく観察して設計するしかなかった。
- 難しい問題2: フレーム問題
 - － 人間が知識を記述することで、人工知能を動作させる。
 - － そのときに、いくら知識を書いても、うまく例外に対応できない。
- 難しい問題3: シンボルグラウンディング問題
 - － シマウマがシマのある馬だと、計算機が理解することができない。
 - － シンボル(記号)がそれが指すものと接続(グラウンド)しておらず、シンボルの操作ができない。

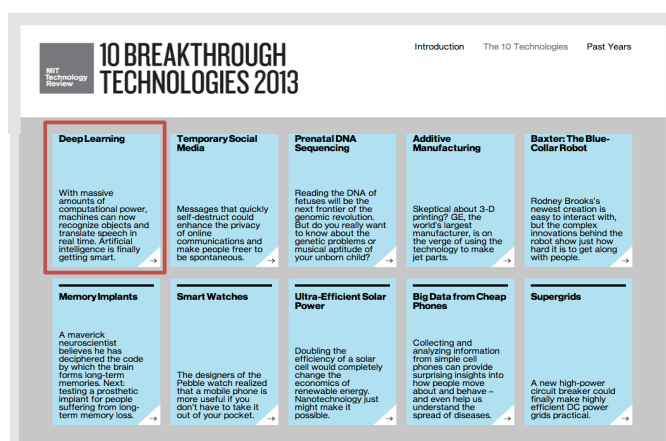


結局のところ、いままでの人工知能は、
人間が現実世界の対象物を観察し、「どこに注目」するかを見ぬいて(特徴量を取り出して)、モデルの構築を行っていた。
その後の処理は自動で行うことができたが、モデル化の部分に人間が大きく介在していた。
それが、唯一にして最大の問題であった。

9

Deep Learning

- AIにおける50年来のブレイクスルー
 - － データをもとに、どこに注目すべきかという「特徴量」が自動的に獲得されている



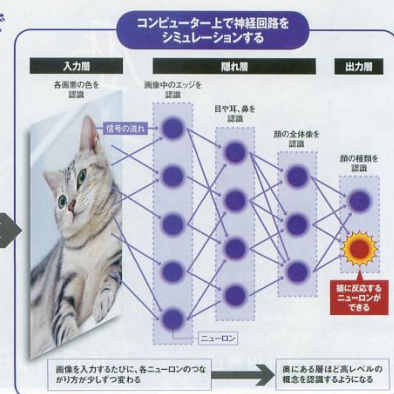
人工の神経回路、威力増す

「ディープラーニング」と呼ぶ人工知能技術が高い関心を集めている。画像や音声の認識精度が大幅に高まるため、米グーグルなどが研究に参入。経済動向の予測や新薬開発などにも威力を発揮する可能性がある。

ここ1～2年、世界中の人工知能の研究から大きな注目を浴びている技術がある。コンピューターに人間と同じように経験に基づいた行動をさせる機械学習の一種で、「ディープラーニング」と呼ばれる新手法だ。インターネット社会を支える画像認識や音声認識、新薬開発に役立つ化合物の活性予測——。こうした技術の精度を競うコンテストで、ディープラーニングが過去の記録を大幅に塗り替え、次々と優勝を果たしている。「これほど飛躍的に精度が向上するとは信じられない」「まさに衝撃的な結果だ」。専門家からは、口々に驚きの声上がる。

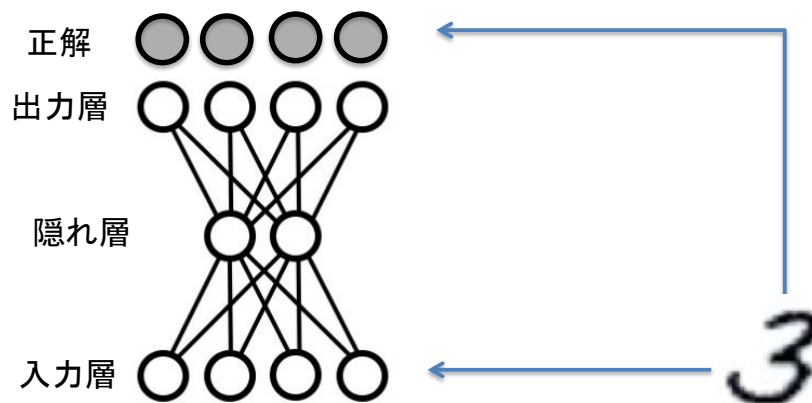
ディープラーニングで画像を認識する流れ

ディープラーニングでは、コンピューター上に人間の脳と似たような多層の神経回路を構築。大量の画像や文字情報を入力してトレーニング(訓練)すると、そこに含まれる高度な概念が自然に引き出される。米グーグルの研究では、出力層のニューロンが、「猫」を認識して強く反応するようになった(右は簡略化した仕組みの図)。



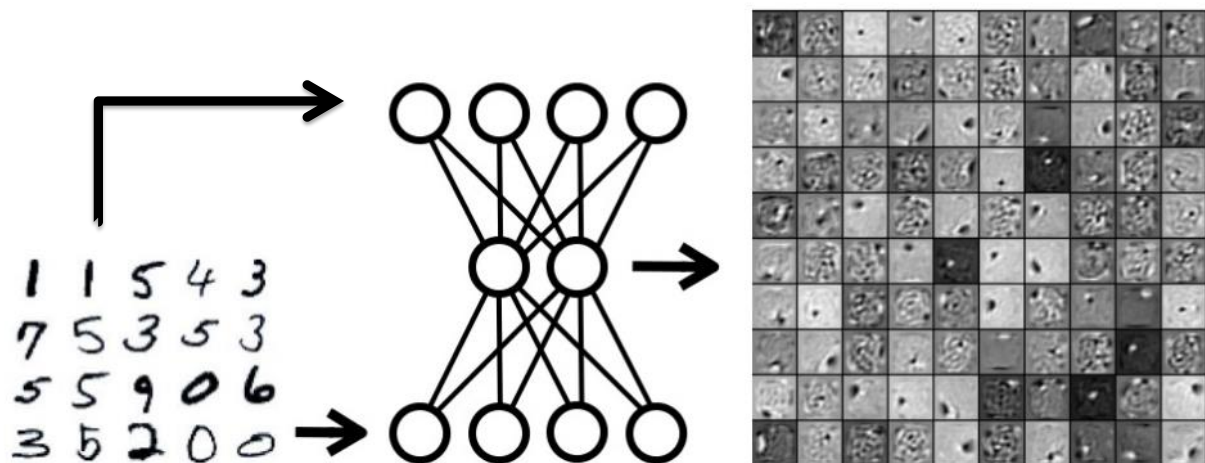
Auto-encoder(2006-)

- Deep Learningの主要な構成要素
- 出力を入力と全く同じにしたニューラルネットワーク
 - 手書き文字認識では、ひとつの画素の値を予測する。
 - 普通に考えると意味ない。
- 「1万円札をお店の人に渡して、1万円札をうけとるようなもの」(「考える脳 考えるコンピュータ」J. Hawkins)
- 隠れ層のノードが「入力を圧縮したもの」になる。



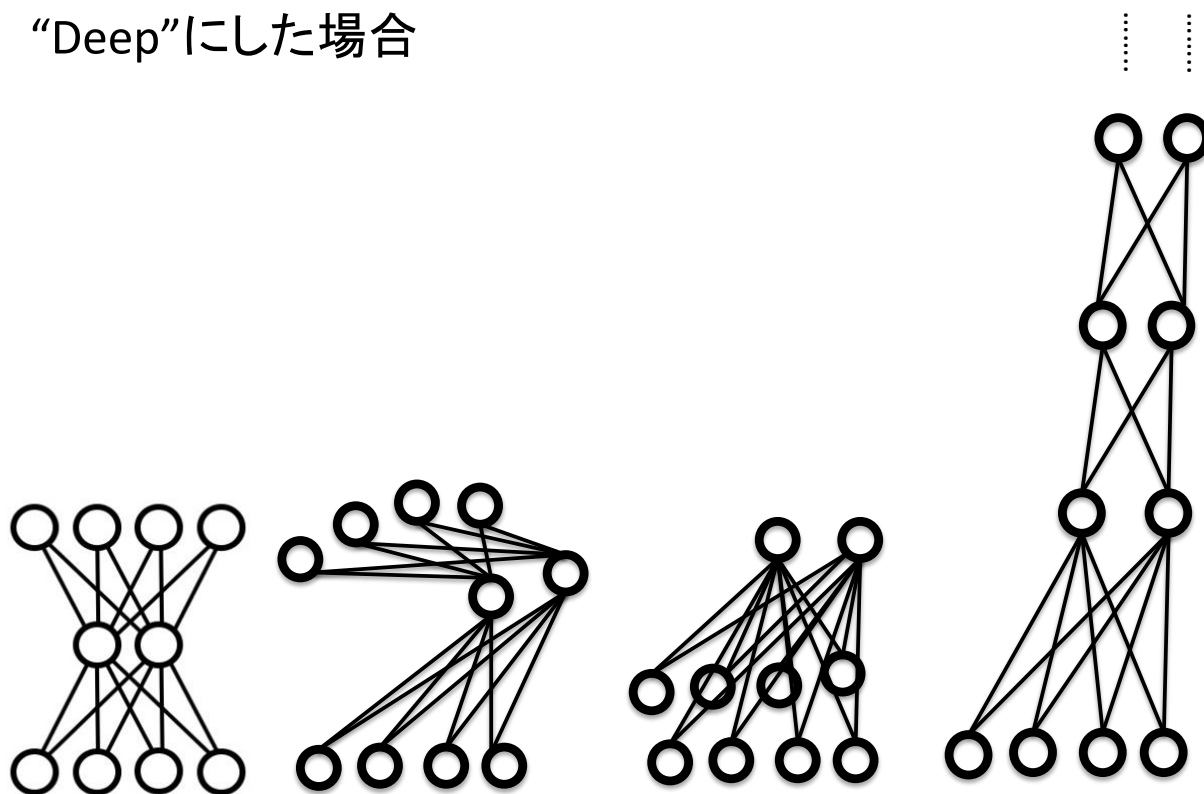
11

Auto-encoderで得られる表現



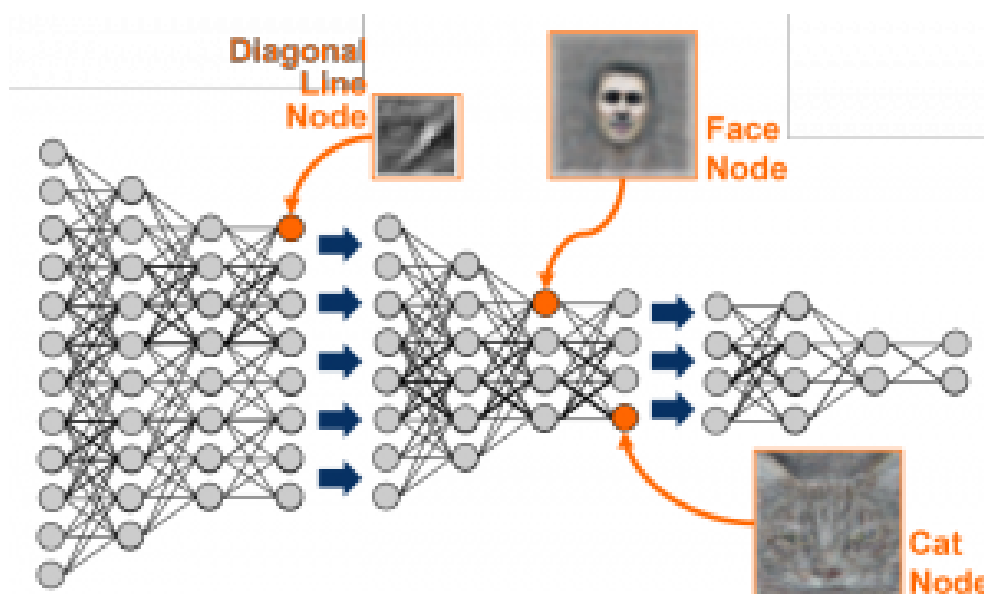
12

“Deep”にした場合



13

Googleの猫(2012)

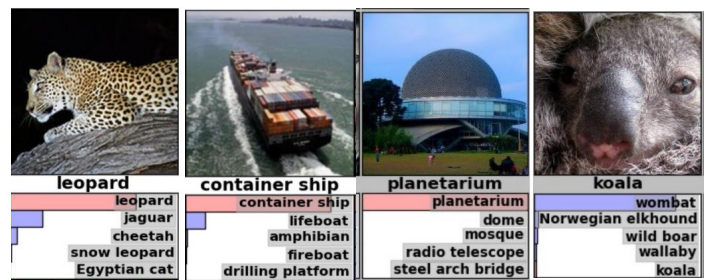


- YouTubeから取ってきた大量の画像をニューラルネットワークに学習させることで、下位の層のニューロンには線や点といった単純な特徴量が、上位の層には、人の顔や猫といったより複雑な特徴量が学習される。
- 人間の視神経のモデルとして知られているものと極めて近い。

14

認識:ディープラーニングの実績(2012)

- ILSVRC2012: Large Scale Visual Recognition Challenge 2012



Team name	Error	Description
SuperVision	15.315%	Using extra training data from ImageNet Fall 2011 release
SuperVision	16.422%	Using only supplied training data
ISI	26.602%	Naïve sum of scores from classifiers using each FC
ISI	26.646%	Naïve sum of scores from classifiers using each FV
ISI	26.952%	Naïve sum of scores from each classifier with SIFT+FV, LBP+FV, GIST+FV and CSIFT+FV, respectively
OXFORD_VGG	26.979%	Mixed selection from High-Level SVM scores and Baseline Scores, decision is performed by looking at the validation performance.
...	...	...

15

認識: 2012年以降のエラー率の変化

	Error
Imagenet 2011 winner (not CNN)	25.7%
Imagenet 2012 winner	16.4% (Krizhevsky et al.)
Imagenet 2013 winner	11.7% (Zeiler/Clarifai)
Imagenet 2014 winner	6.7% (GoogLeNet)
Baidu Arxiv paper:2015/1/3	6.0%
Human: Andrej Karpathy	5.1%
Microsoft Research Arxiv paper: 2015/2/6	4.9%
Google Arxiv paper: 2015/3/2	4.8%
Microsoft Research CVPR paper: 2015/12/10	3.6%
Latest	3.1%

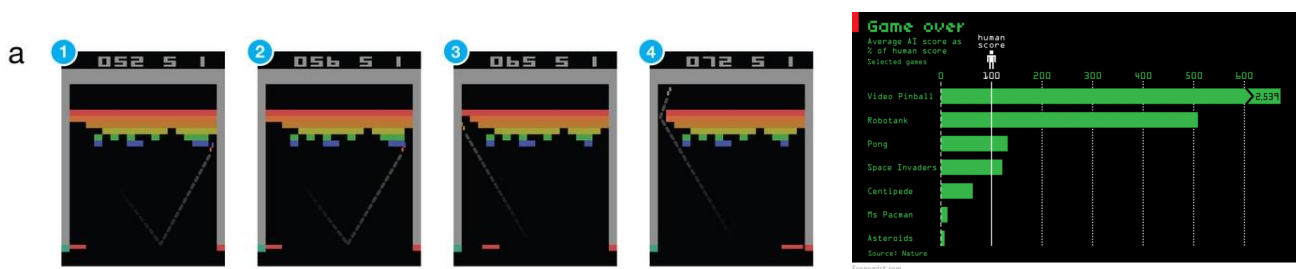
2015年2月には人間の精度を超えた

画像認識で人間の精度を超えることは数十年間、実現されていなかった

16

運動の習熟: ディープラーニング+強化学習(2013-)

- 強化学習とは、行動を学習する仕組み。
 - 「報酬」が得られると、事前の行動を強化する。
 - 「状態」「行動」→「望ましき(報酬ありなし)」
 - 古くからある技術だが、これまでは、「状態」を人間が定義してきた。
- 運動の習熟が可能に
 - 状態の認識に、ディープラーニングを使う。
 - DeepMindの研究者(D. Hassabisら)。その後、Googleが買収。
- 試行錯誤することによって、運動が習熟する
 - 最初は下手。繰り返すうちに、うまくなっていく。
 - 最終的には、ブロック崩しでの通路を作ったり、インベーダーゲームでの「名古屋撃ち」も。
 - 「全く同じプログラム」で、異なるゲームを学習。半数のゲームで人間のハイスコアを上回る

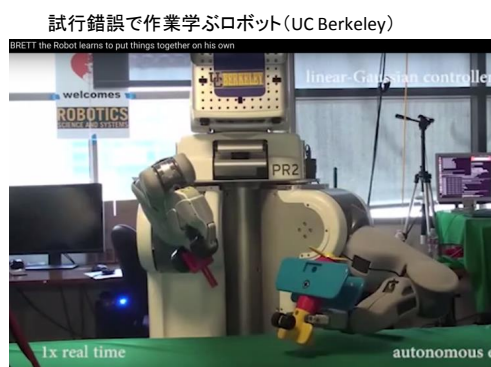


<http://www.clubic.com/mag/actualite-756059-google-jeu-video.html>
<http://www.economist.com/news/briefing/21650526-artificial-intelligence-scares-people-excessively-so-rise-machines>

17

運動の習熟: ディープラーニング+強化学習が実世界へ(2015-)

- 実世界への適用
 - 2015年5月 試行錯誤で部品の取付を習熟するロボットの開発(UC Berkeley)
 - 2015年5月 試行錯誤で運転を習熟するミニカーの開発(PFN社, 日本)
 - 2015年12月 試行錯誤でピッキングが上達するロボットの開発(PFN・ファナック, 日本)
 - その他、メリーランド大、EUのプロジェクト等も進展
- 考えてみれば当たり前
 - 犬や猫でもできる。高次の言語能力は必要ない。認識が問題だった。
 - 歴史的には、多数の人工知能研究者がこのことを主張してきた。



<https://research.preferred.jp/2015/06/distributed-deep-reinforcement-learning/>
<http://news.berkeley.edu/2015/05/21/deep-learning-robot-masters-skills-via-trial-and-error/>
http://www.nikkei.com/news/print-article/7R_FLG=0&bf=0&ng=DGXMZ08384452055A300C1000000&uah=DF170520127709



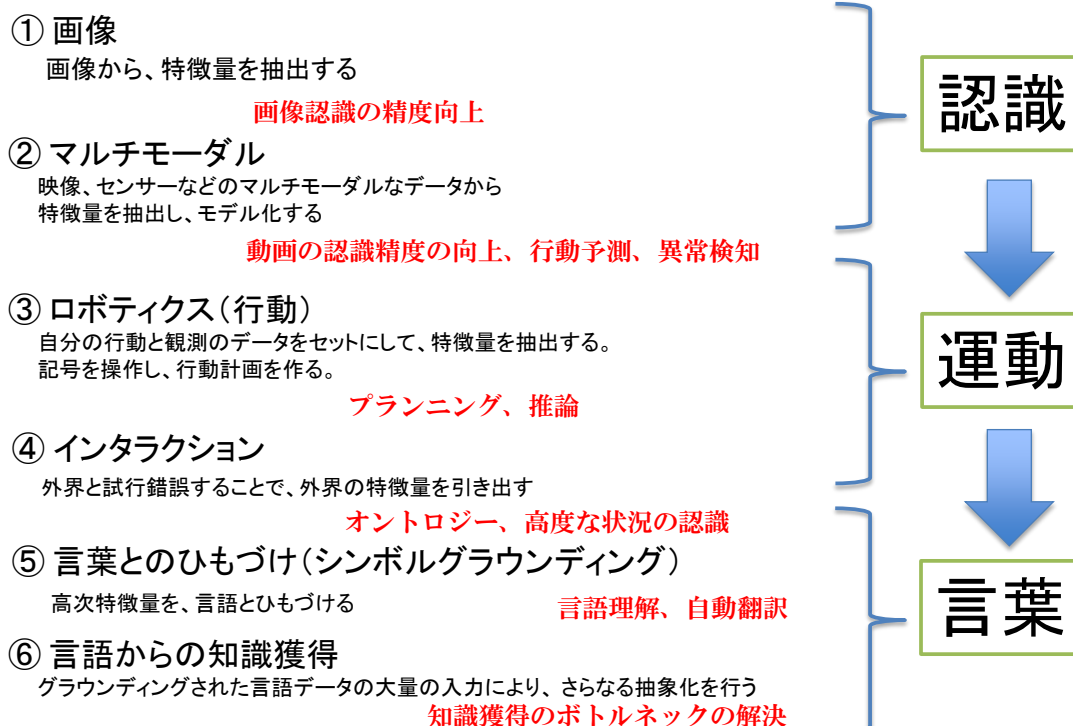
18

ディープラーニングの人工知能における意味

- モラベックのパラドックス:「子供のできることほど難しい。」
 - 高度な推論よりも、認識や運動スキルの方が難しい。
 - それがここ3年くらいのあいだに一気にできるようになった
- 現在のコンピュータのパワーでようやく可能に
 - GPUを数十台並列に並べて、数日～数ヶ月計算させてようやく精度が上がる
- アイディアは昔からあった。もともとは日本発
 - 1980年当時、NHK放送技術研究所にいた福島邦彦先生によるネオコグニトロン
 - その後も多くの研究者が試みている
- 初期仮説への回帰
 - 初期仮説「なぜ知能をコンピュータで実現することはできないのか？」
 - できると思っていた→できない理由があった→それが解消された→だとしたら、もう一度できるという仮説を取るべきでは。
 - 産業として非常に大きい可能性を秘めている。

19

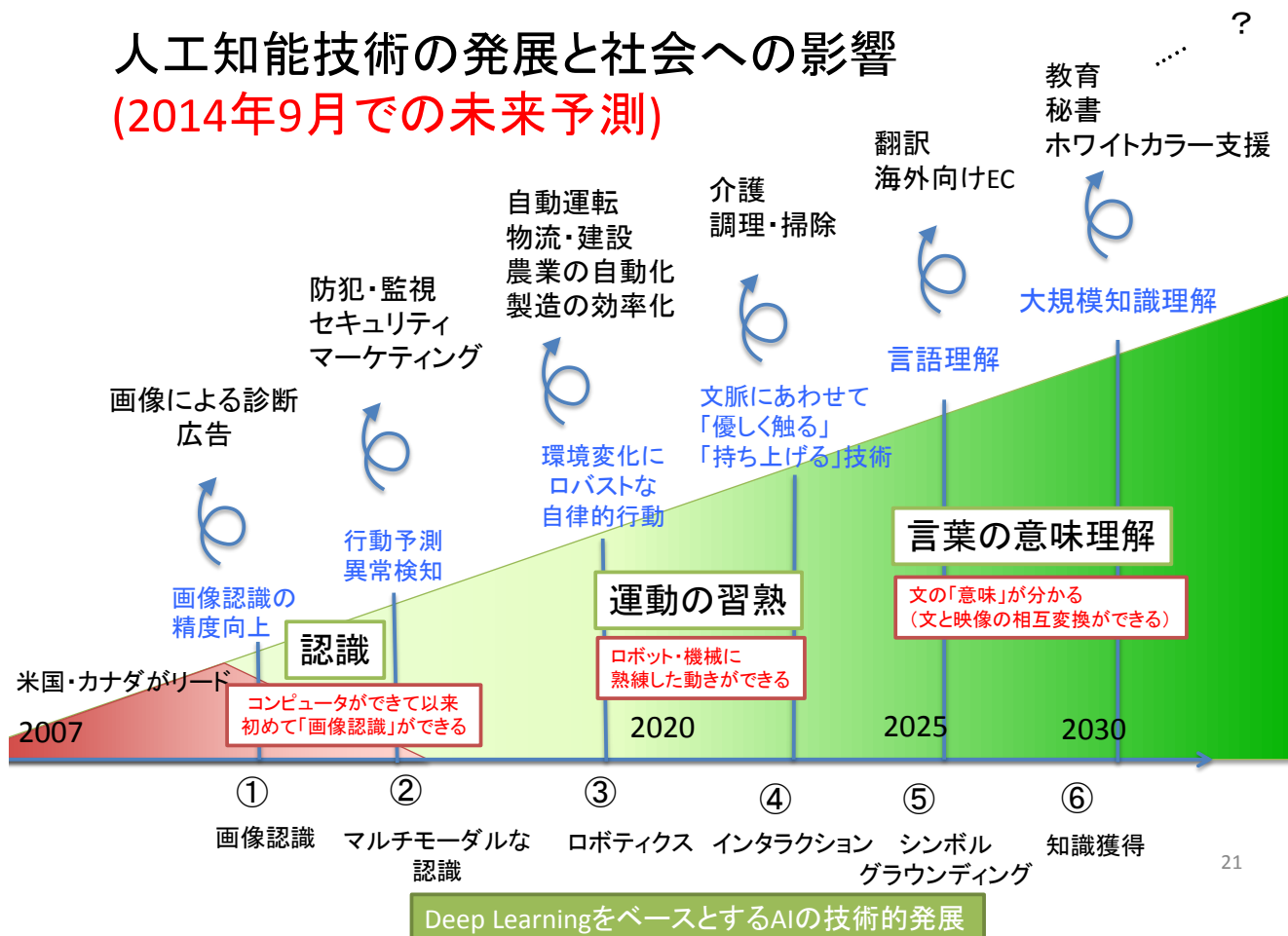
ディープラーニングの今後の発展



ディープラーニングがすごいというより
その先に広がる世界がすごい

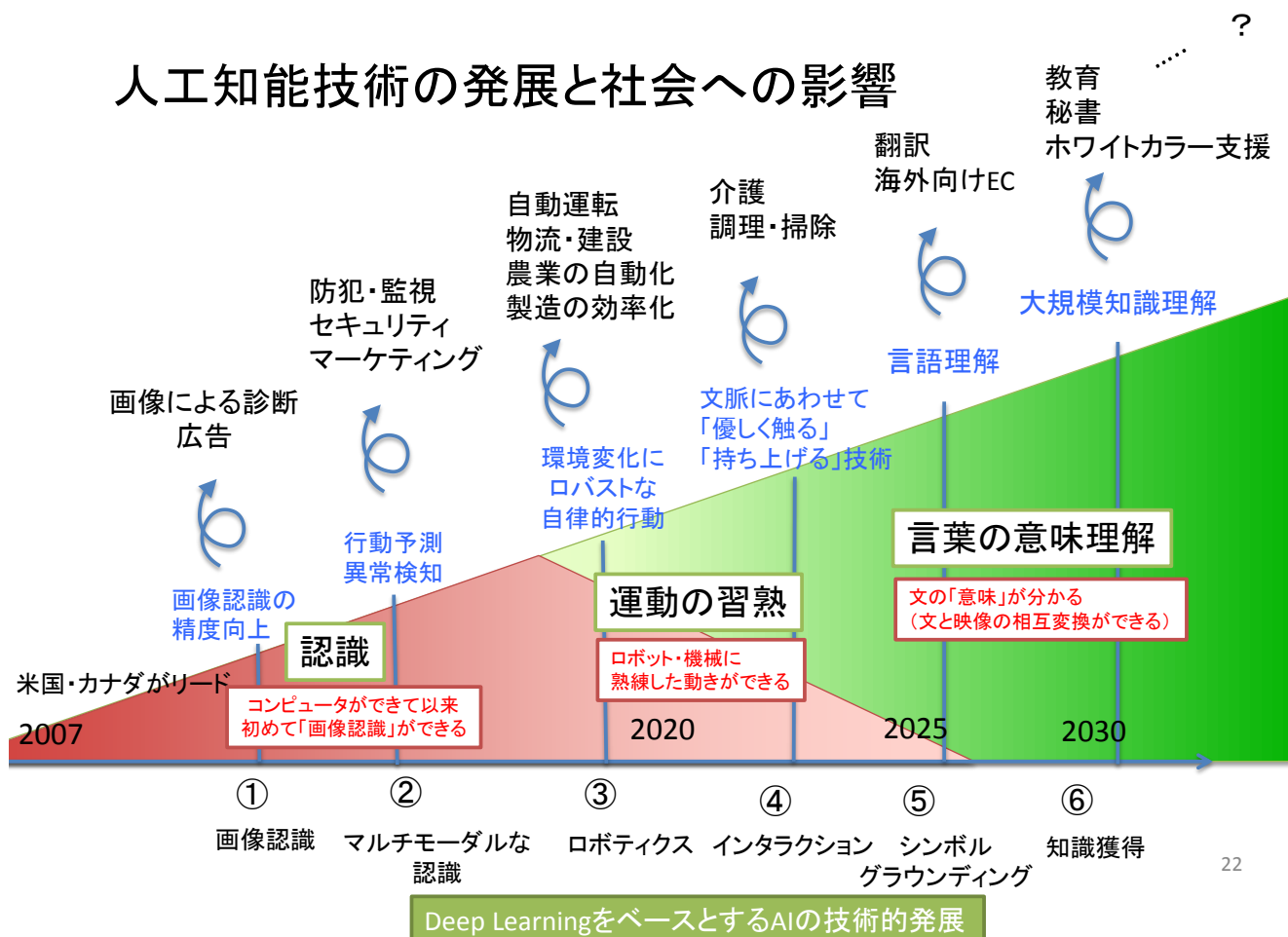
20

人工知能技術の発展と社会への影響 (2014年9月での未来予測)



21

人工知能技術の発展と社会への影響

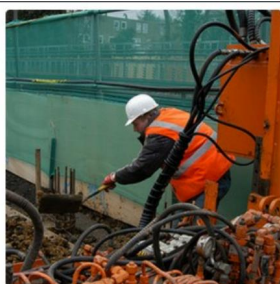


22

言葉の意味理解: Automated Image Captioning (2014-)



"man in black shirt is playing guitar."



"construction worker in orange safety vest is working on road."



"two young girls are playing with lego toy."



"boy is doing backflip on wakeboard."



"girl in pink dress is jumping in air."



"black and white dog jumps over bar."



"young girl in pink shirt is swinging on swing."

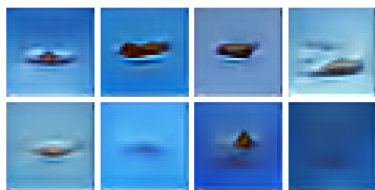


"man in blue wetsuit is surfing on wave."

<http://cs.stanford.edu/people/karpathy/sfmltalk.pdf>

23

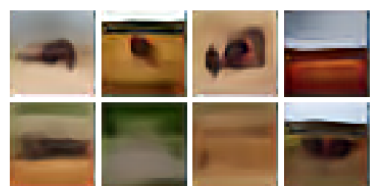
言葉の意味理解: Generating Images (2015.12-)



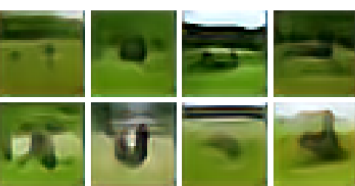
A very large commercial plane flying in blue skies.



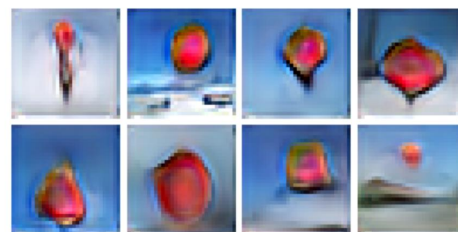
A very large commercial plane flying in rainy skies.



A herd of elephants walking across a dry grass field.



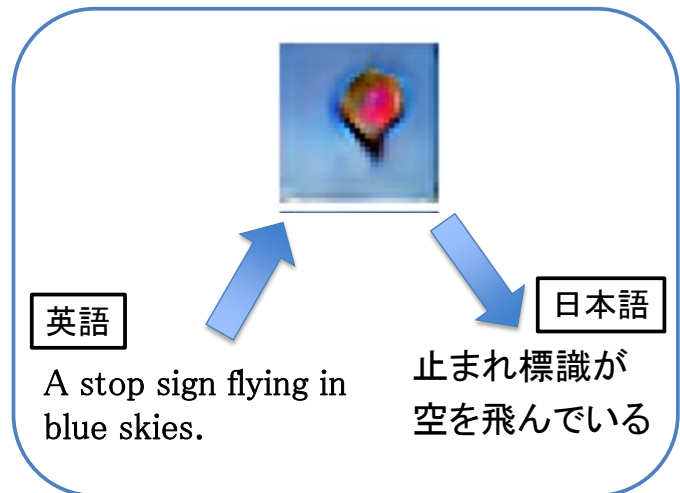
A herd of elephants walking across a green grass field.



A stop sign flying in blue skies.

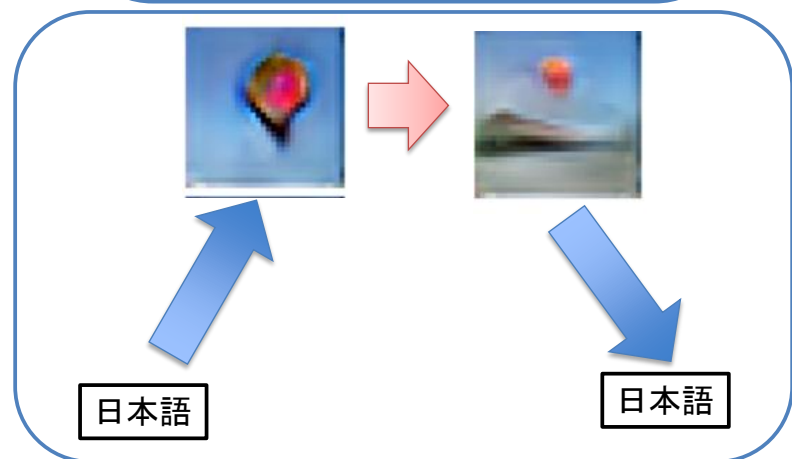
- 画像による翻訳(意識)

- 日本語→画像→英語
- 課題
 - 画像から映像、体験
 - 抽象概念の扱い
 - 人間固有のセンサー系
- いずれにしても視覚的な機構がベースにあるのは間違いない



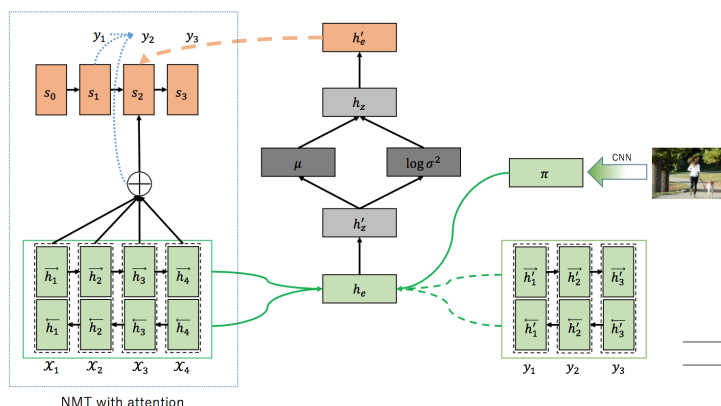
- 映像による推論

- 言語→映像→(シーン予測)→次の映像→言語
- 「風船が飛んでいる」→「山まで飛んで行くのかな」



画像を介した新しい翻訳方式の提案(松尾研)

- 文から画像を生成し、もとの文と生成した文から翻訳文を生成する。
- 新しいモデルを提案
- 従来手法より制度が向上

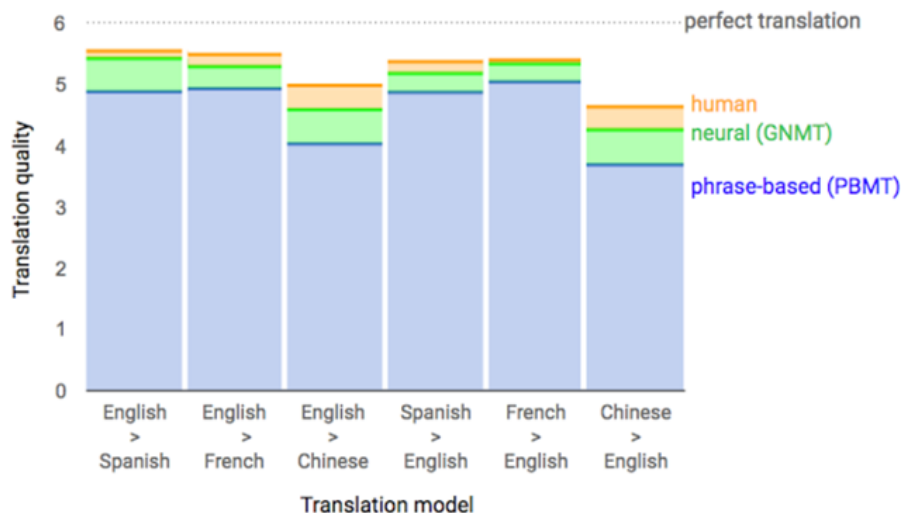


ソース	a woman does a somersault on a trampoline on the beach.
正解	eine frau macht einen salto auf einem trampolin am strand.
VNMT	eine frau macht einen salto am strand.
提案モデル	eine frau macht einen salto auf einem trampolin am strand.

	METEOR ↑		BLEU ↑	
	val	test	val	test
NMT	50.06 (53.01)	49.67 (53.72)	33.1	33.9
VNMT	50.29 (53.09)	49.66 (53.57)	33.4	33.7
CMU	- (-)	50.8 (54.1)	-	35.1
N	50.50 (53.70)	50.44 (54.43)	34.5	34.6
Our Model	51.26 (54.93)	51.56 (55.60)	35.1	35.1

Google翻訳がDL方式に(2016.9)

- Google Neural Machine Translation (GNMT)
- 全部入り:8層のbidirectional RNN (LSTM), Residual connection, Attention
- さらなる工夫:TPUの利用, Wordpiece
- 60%エラーを減らした。
- “two to three decimal orders of magnitudes bigger than the WMT corpora for a given language pair”



<https://research.googleblog.com/2016/09/a-neural-network-for-machine.html>

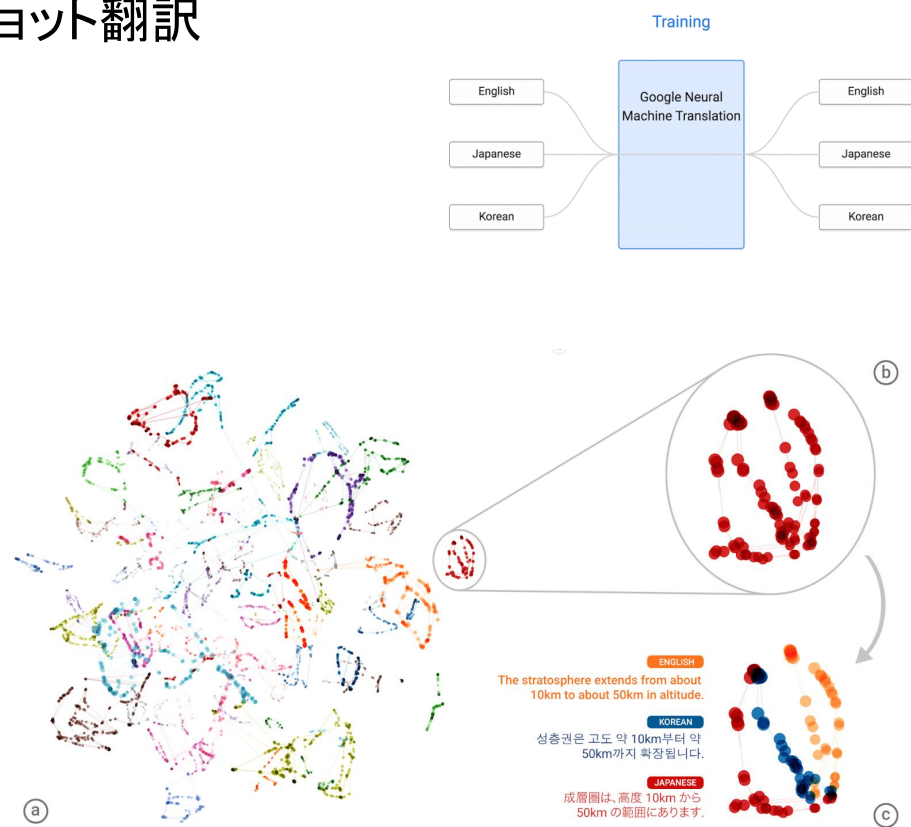
27

Deep reinforcement learning (deep RL) has been successful in learning sophisticated behaviors automatically; however, the learning process requires a huge number of trials. In contrast, animals can learn new tasks in just a few trials, benefiting from their prior knowledge about the world. This paper seeks to bridge this gap. Rather than designing a “fast” reinforcement learning algorithm, we propose to represent it as a recurrent neural network (RNN) and learn it from data. In our proposed method, RL2, the algorithm is encoded in the weights of the RNN, which are learned slowly through a general-purpose (“slow”) RL algorithm. The RNN receives all information a typical RL algorithm would receive, including observations, actions, rewards, and termination flags; and it retains its state across episodes in a given Markov Decision Process (MDP). The activations of the RNN store the state of the “fast” RL algorithm on the current (previously unseen) MDP. We evaluate RL2 experimentally on both small-scale and large-scale problems. On the small-scale side, we train it to solve randomly generated multi-armed bandit problems and finite MDPs. After RL2 is trained, its performance on new MDPs is close to human-designed algorithms with optimality guarantees. On the large-scale side, we test RL2 on a vision-based navigation task and show that it scales up to high-dimensional problems.

深い強化学習(深いRL)は、洗練された行動を自動的に学習するのに成功しています。しかし、学習プロセスには膨大な試行が必要です。これとは対照的に、動物は世界についての以前の知識から恩恵を受け、わずかな試行で新しい仕事を習得することができます。このペーパーは、このギャップを埋めようとしています。「高速」強化学習アルゴリズムを設計するのではなく、それをリカレントニューラルネットワーク(RNN)として表現し、それをデータから学習することを提案します。提案された方法RL2では、アルゴリズムはRNNの重みに符号化され、RNNは汎用(「遅い」)RLアルゴリズムによってゆっくり学習される。RNNは、観察、行動、報酬、および終了フラグを含む典型的なRLアルゴリズムが受け取るすべての情報を受信する。それは所与のマルコフ決定プロセス(MDP)においてエピソード全体にわたってその状態を保持する。RNNの活性化は、現在の(以前は見えなかった)MDP上の「高速」RLアルゴリズムの状態を記憶する。我々は、RL2を小規模問題と大規模問題の両方について実験的に評価する。小規模な面では、無作為に生成された複数武装の禁止問題と有限のMDPを解決するように訓練します。RL2が訓練された後、新しいMDPでの性能は、最適性が保証された人間が設計したアルゴリズムに近くなります。大規模な側面では、RL2をビジョンベースのナビゲーションタスクでテストし、高次元の問題までスケールアップすることを示します。

28

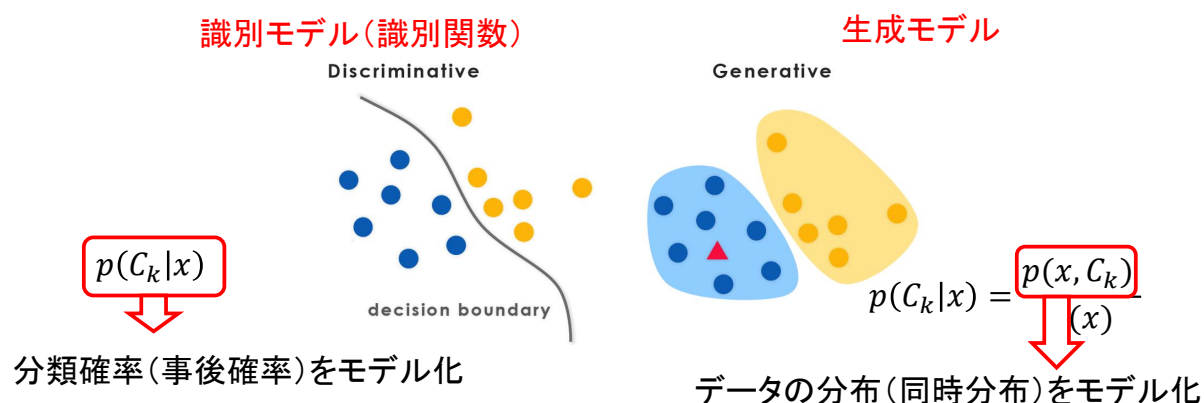
ゼロショット翻訳



<https://research.googleblog.com/2016/11/zero-shot-translation-with-googles.html>

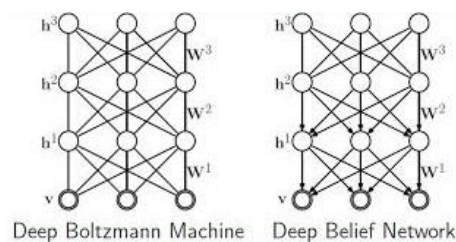
29

鍵となるのは、深層生成モデル



- データを分けることのみに興味があるのが識別モデル
 - 深層NNやSVMなどは識別モデル(正確には識別関数)
- データの生成源も考えるのが生成モデル
- 識別モデルの方が圧倒的に学習は簡単

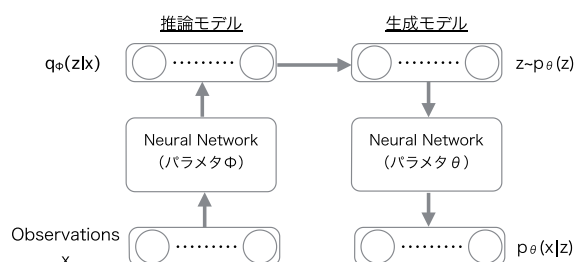
深層生成モデル



- 深層生成モデル (deep generative model)
 - 潜在変数が多層になったモデル
 - より複雑なモデルを学習することができる。
- 以前から研究されていた。
 - Deep belief network (stacked RBM) [Hinton+ 2006]
 - Deep Learningの元祖
 - 最終層以外は有向グラフ, 最終層は無向グラフのグラフィカルモデル
 - Deep Boltzmann machine [Salakhutdinov 2009]
 - 全ての層が無向グラフのグラフィカルモデル
- その後着目されなくなったが、2014年頃から、再び着目されるようになった。
 - それに合わせて教師なし学習も復権。
- 最近の深層生成モデルは次の2つが代表的
 - Variational autoencoder (VAE)
 - Generative adversarial network (GAN)

Variational autoencoder (VAE)

- モデル



- 学習: 変分下界を最大化

$$\hat{\mathcal{L}}(\theta, \phi; \mathbf{x}) = -D_{KL}(q_{\phi}(\mathbf{z}|\mathbf{x})||p_{\theta}(\mathbf{z})) + \frac{1}{L} \sum_{l=1}^L \log p(\mathbf{x}^{(l)}|\mathbf{z}^{(l)}, \theta)$$

$$\text{ただし } \mathbf{z}^{(l)} = g(\epsilon^{(l)}, \mathbf{x}), \epsilon^{(l)} \sim p(\epsilon)$$

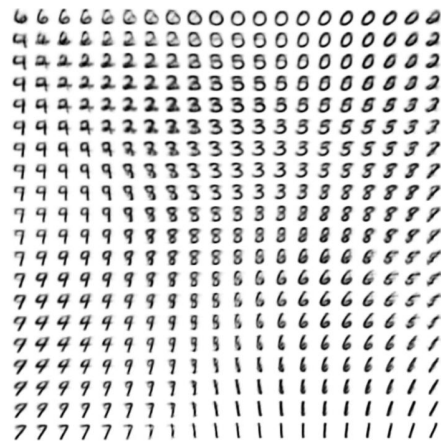
- 第1項が正規化項、第2項を負の再構成誤差

VAEによる生成画像

- 数字の生成

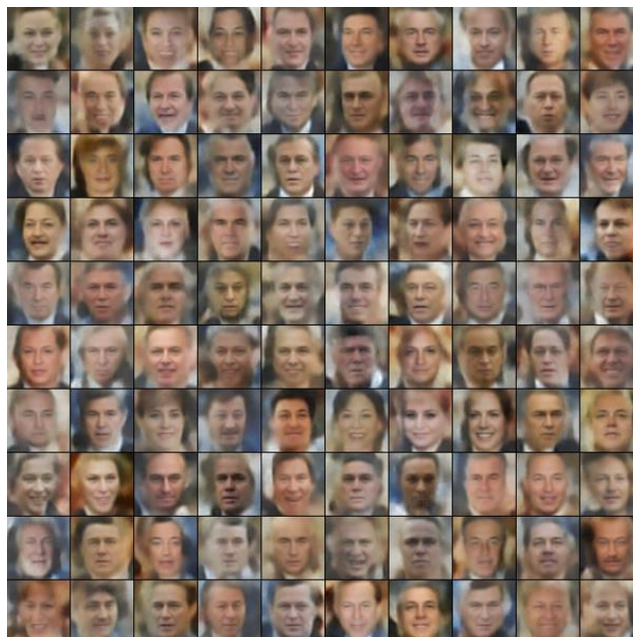


- 潜在空間の可視化
 - 多様体学習ができています.



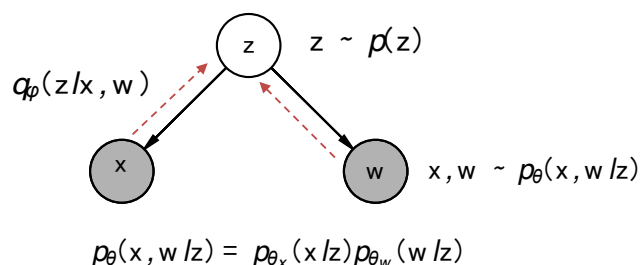
VAEによる生成画像

- 顔の生成(ネットワークにCNNを利用)



マルチモーダル変分オートエンコーダ(松尾研)

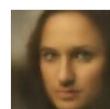
2つのモーダル x と w の同時確率 $p(x, w)$ を求める新しいモデルを提案



	$\leq \log p(x)$			$\leq \log p(x w)$	
	multiple	single		multiple	single
VAE-GAN		-4439	CVAE-GAN		-4152
JMVAE-GAN	-4141	-4144	CMMA-GAN		-4147
			JMVAE-GAN		-4130



元画像



再構成画像



Smiling=3にした再構成画像

$$\mathcal{L}(x, w) = -D_{KL}[q_\phi(z|x, w) || p(z)] + E_{q_\phi(z|x, w)} [\log p_{\theta_x}(x|z) + \log p_{\theta_w}(w|z)]$$

エンコーダ $q_\phi(z|x, w)$ 、デコーダ $p_{\theta_x}(x|z)$, $p_{\theta_w}(w|z)$

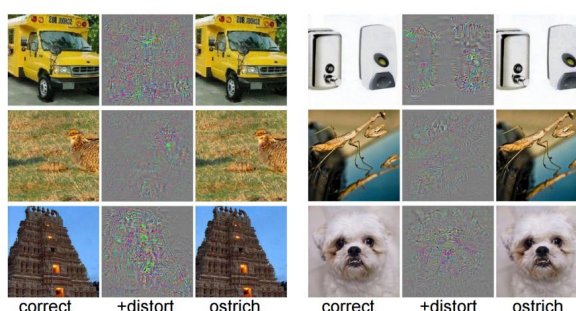
- ほかにも、密度比推定(GAN) Generative Adversarial Network

Convolutional Neural Network (CNN) は騙されやすい

Source: [Szegedy et al. 2013]

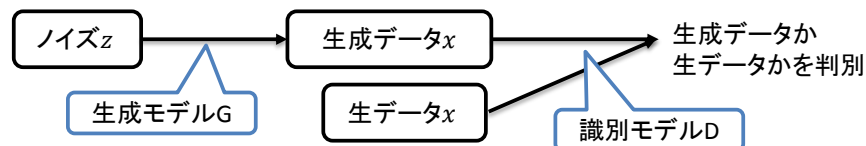


Bus + Noise = Ostrich



Generative adversarial network (GAN)

- Generative adversarial network [Goodfellow+ 2014]
 - 敵対的生成ネットワーク
 - 生成モデルG: 潜在変数(ノイズ) z から x を生成する写像
 - 識別モデルD: サンプル x が生成されたものか生データかを判別



- 学習: GとDで次のゲーム(ミニマックス法)をする.
 - GはなるべくDを騙すようにする
 - Dはなるべくうまく判別できるようにする.

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

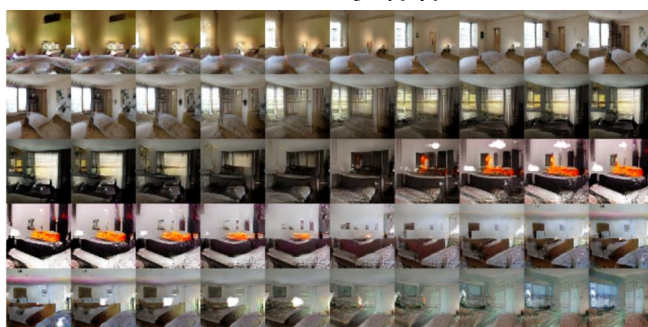
→生データと同じ x が生成できるようになる

GAN系の手法の例

すべて生成したbedroom(DCGAN)



bedroom多様体



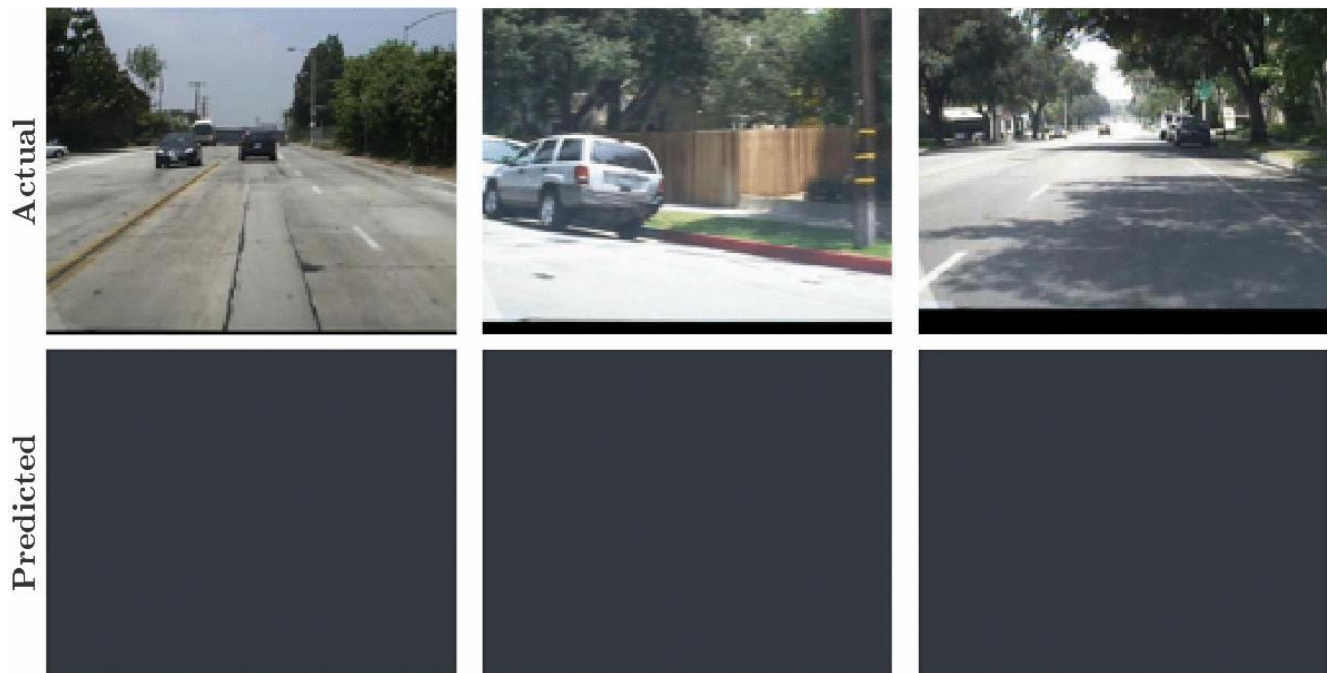
本物 生成 生成



Generative ConvNet

どんどん本物の画像に近づきつつある

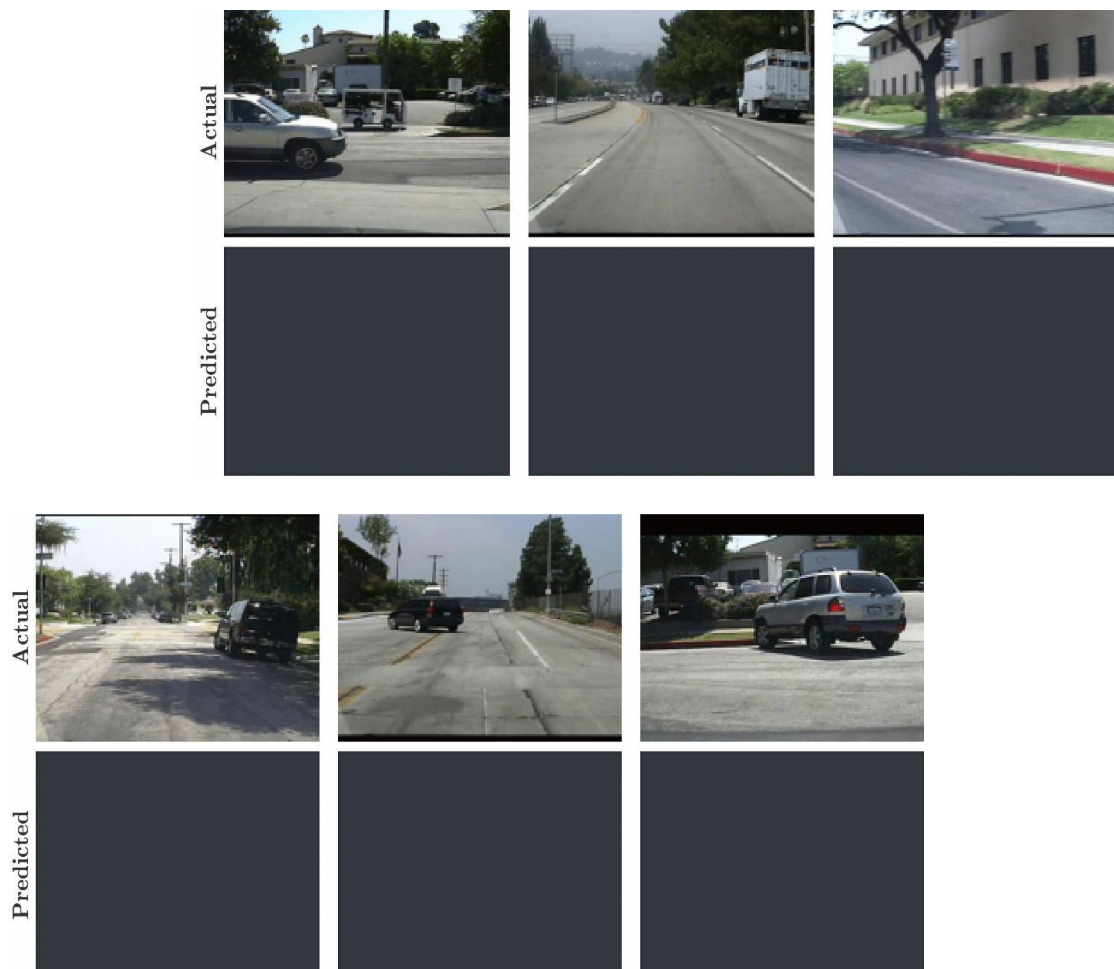
未来の画像を予測する



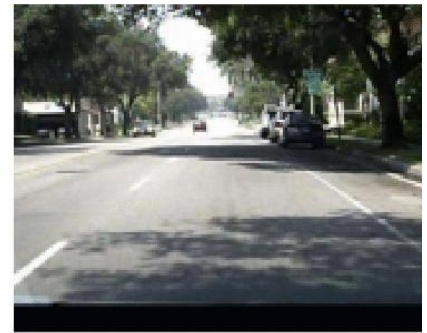
- KITTIデータセット。41,000のフレームから学習。
- 次のフレームを予測する。10フレーム=1秒。4レイヤーのモデル。

41

William Lotter, Gabriel Kreiman, David Cox, Deep Predictive Coding Networks for Video Prediction and Unsupervised Learning, CoRR abs/1605.081042016, 2016



42



- より先の未来を予測。5フレーム(0.5秒後)になると、ややぼやけてくる。

43

静止画だけを与えて、1秒後を予測する

- ラベルなしの動画から、動画認識(行動分類)と動画生成(未来予測)の両方に使えるシーンダイナミクスを学習する。
- CNNによる時空間の畳み込み+GANを使って、背景と前景を切り分ける。
- 1秒までの短いビデオをフルフレームで生成できる。

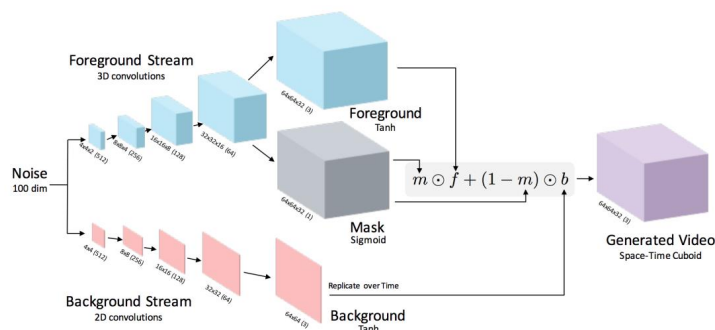


Figure 1: **Video Generator Network:** We illustrate our network architecture for the generator. The input is 100 dimensional (Gaussian noise). There are two independent streams: a moving foreground pathway of fractionally-strided spatio-temporal convolutions, both of which up-to create the generated video using a mask from the motion and the number of channels in parenthesis.

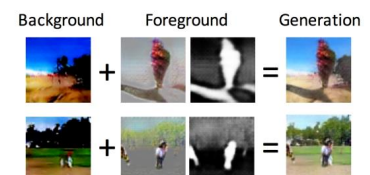
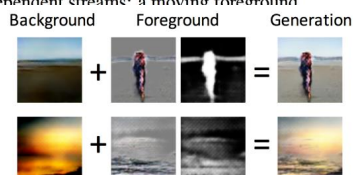
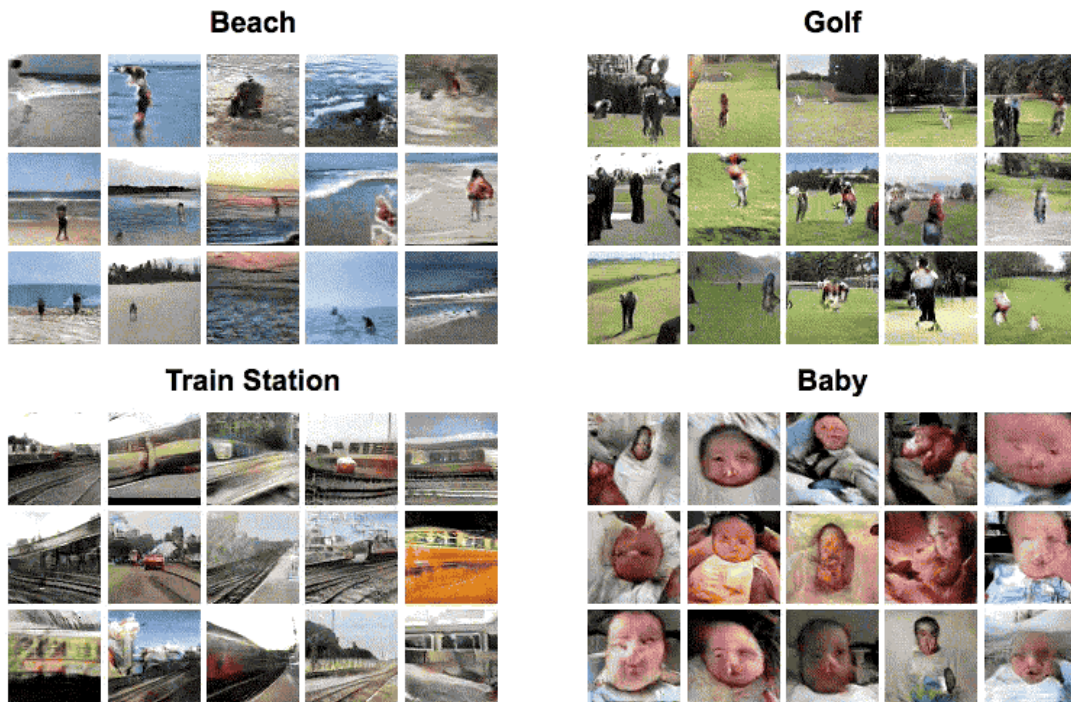


Figure 3: **Streams:** We visualize the background, foreground, and masks for beaches (left) and golf (right). The network generally learns to disentangle the foreground from the background.



- 最初のフレームだけ与え、次の1秒を生成することができる。
- 200万の動画 (Flickr) から5000時間分のデータで学習。

45

動作とその帰結の予測



- Unsupervised Learning for Physical Interaction through Video Prediction (2016)
- 5万本のロボットのインタラクション(押す動作)の動画でも学習。
- CNNに、ロボットのアクションを途中で入れて、見える画像を予測する。

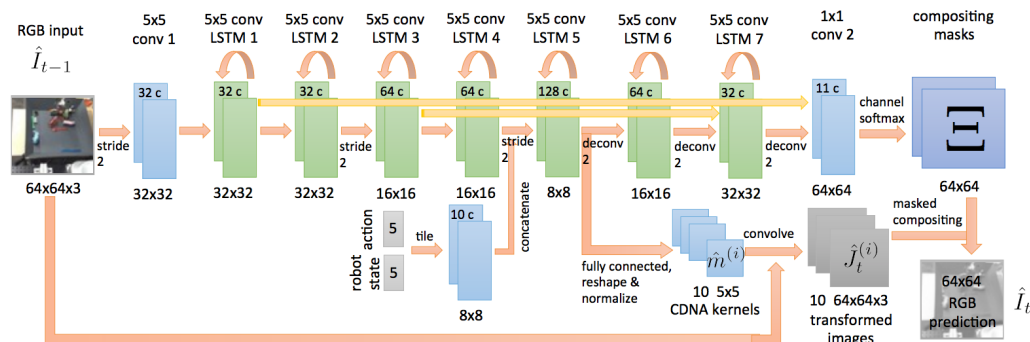


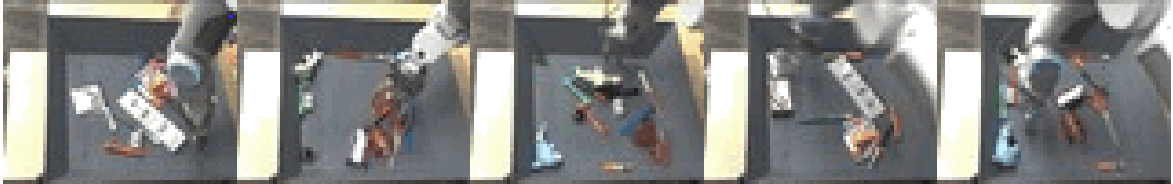
Figure 1: Architecture of the CDNA model, one of the three proposed pixel advection models. We use convolutional LSTMs to process the image, outputting 10 normalized transformation kernels from the smallest middle layer of the network and an 11-channel compositing mask from the last layer (including 1 channel for static background). The kernels are applied to transform the previous image into 10 different transformed images, which are then composited according to the masks. The

46

True



Predicted



True



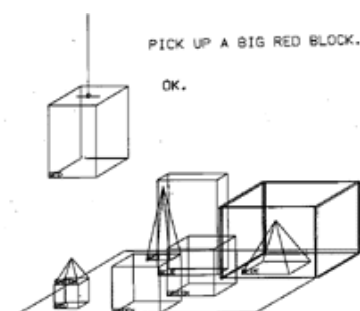
Predicted



SHRDLU (Terry Winograd 1968)

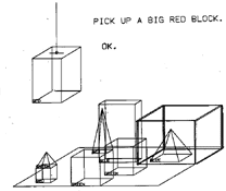
- 1968 年から 1970 年にかけて Terry Winograd が開発したシステム
- 画面の中の「積み木の世界」に、ブロックや円錐、球などが存在し、ユーザからのさまざまな質問に自然言語文で答えることができた。
 - 例えば、「円錐は何に支えられているか？」
- 自然言語文の命令により動かすことができた。
 - 例えば、ユーザは、「緑色の円錐を赤いブロックの上に置け」と指示した後、「その円錐を取り除け」と指示することができた。

Person: Pick up a big red block.
 Computer: OK.
 Person: Grasp the pyramid.
 Computer: I DON'T UNDERSTAND WHICH PYRAMID YOU MEAN.
 Person (changing their mind): Find a block which is taller than the one you are holding and put it into the
 Computer: BY "IT", I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER THAN THE ONE I AM HOLDING
 Computer: OK.
 Person: What does the box contain?
 Computer: THE BLUE PYRAMID AND THE BLUE BLOCK.
 Person: What is the pyramid supported by?
 Computer: THE BOX.
 Person: How many blocks are not in the box?
 Computer: FOUR OF THEM.
 Person: Is at least one of them narrower than the one which i told you to pick up?
 Computer: YES, THE RED CUBE.
 Person: Is it supported?



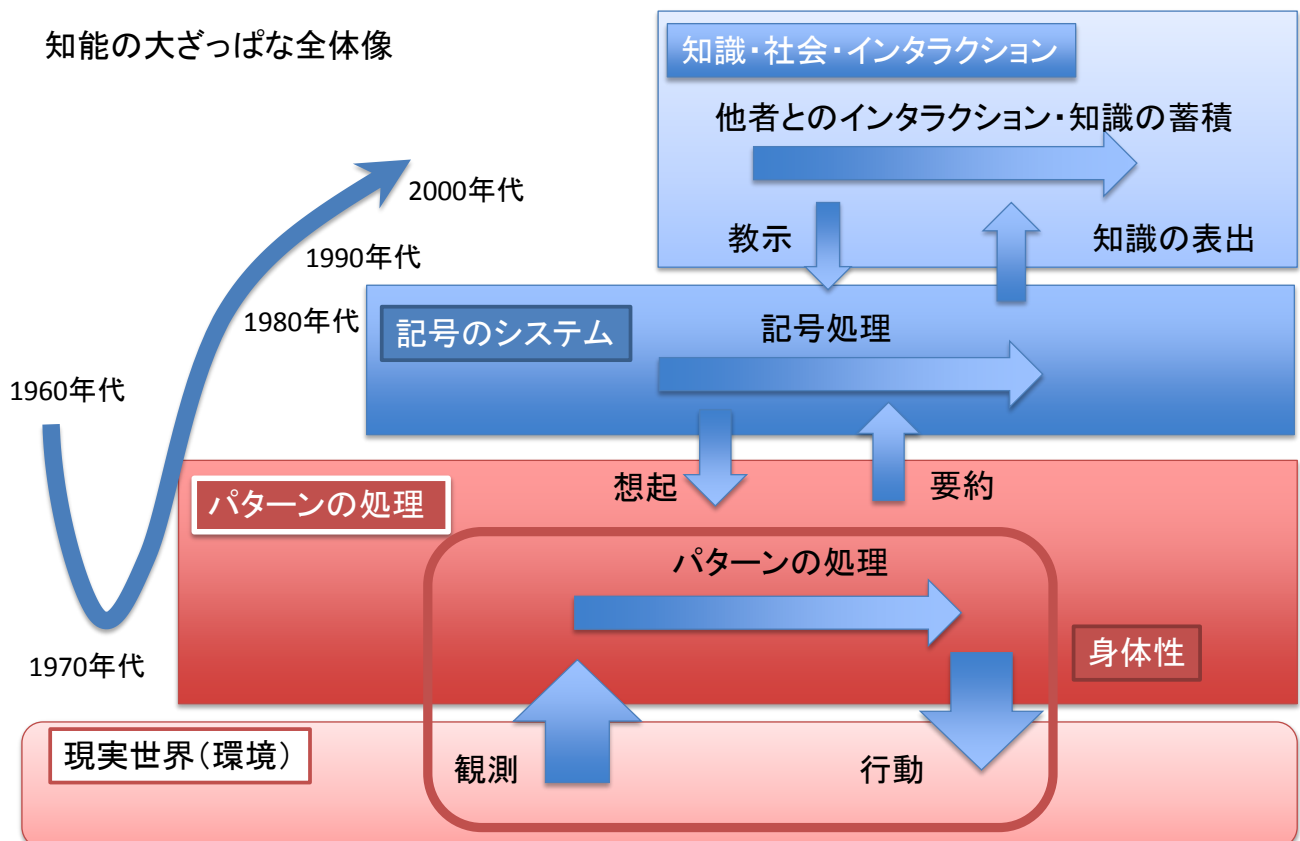
AIのこれまでの研究

Person: Pick up a big red block.
 Computer: OK.
 Person: Grasp the pyramid.
 Computer: I DON'T UNDERSTAND WHICH PYRAMID YOU MEAN.
 Person (changing their mind): Find a block which is taller than the one you are holding and put
 Computer: BY TT, I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER THAN THE ONE I AM
 Holding.
 Person: OK.



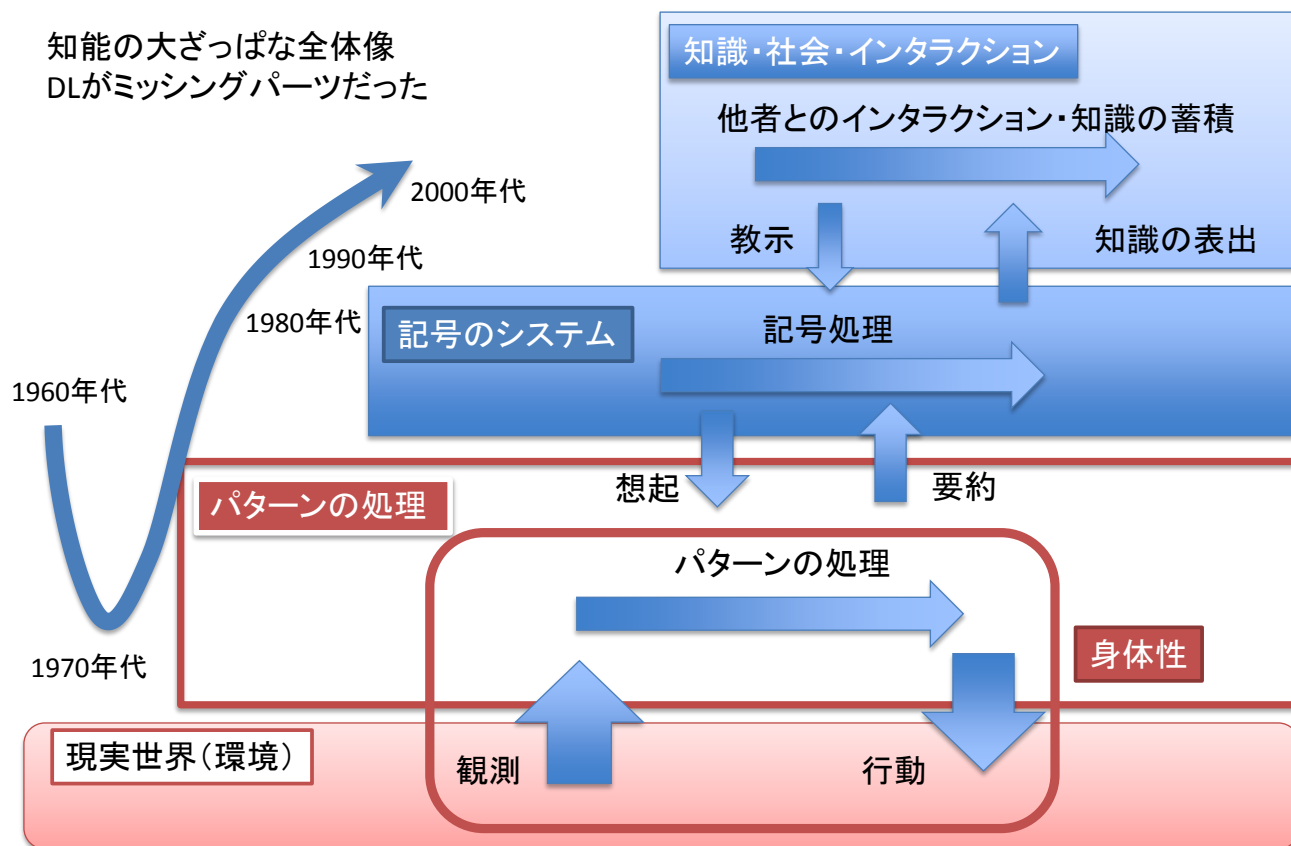
- Terry Winograd :
 - SHRDLU (1968)
 - 画面の中の「積み木の世界」に、ブロックや円錐、球などが存在し、ユーザからのさまざまな質問に自然言語文で答えたり、命令を受けることができた。
 - 例えば、「円錐は何に支えられている か?」「緑色の円錐を赤いブロックの 上に置け」「その円錐を取り除け」など。
- Rodney Brooks
 - 表象なき知能 : Intelligence without representation (1987)
 - 知的能力は、「生存と生殖を最低限保持するのに十分なほど周囲を知覚し、動的な環境世界を動き回ることのできる能力」を土台として、その上に築かれるべきだとした。そして、昆虫型ロボット等によって、環境とインタラクションするだけで、十分に知的に見える振る舞いが出現することを示した。
- Marvin Minsky
 - 象はチェスをしない : Elephants don't play chess (1990)
 - Marvin Minsky は、Brooks の表象なき知能という主張を痛烈に批判した。動物程度の知能は実現できても、人間の知能にとって 記号操作は必要不可欠だという立場であった。
- 環境世界を知覚する仕組みの上に、世界を予測する生成モデルが築かれ、その上にさらに記号の操作が実現されるということではないだろうか。

49



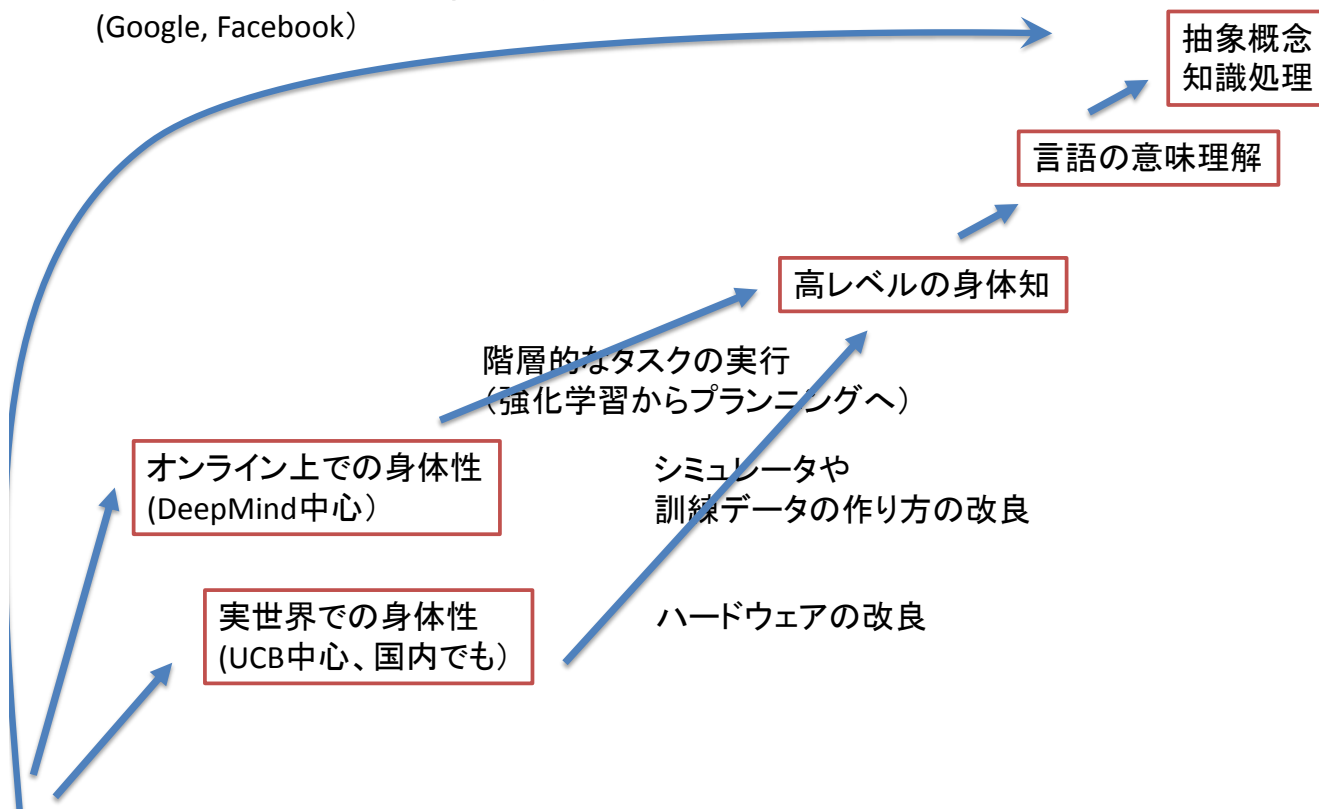
- いままでの推論(述語・命題論理による演繹や帰納・仮説推論)は、記号の空間だけでやろうとしてきた。
- 思考とは、パターンの空間と記号の空間をいったりきたりすること。この上に、言語によるコミュニケーションや知識の蓄積が構成される。
- いずれも目的は、「いかに少ないサンプルで自由度の高いモデルを同定するか」

知能の大ざっぱな全体像
DLがミッシングパーツだった



- ・ いままでの推論(述語・命題論理による演繹や帰納・仮説推論)は、記号の空間だけでやろうとしてきた。
- ・ 思考とは、パターンの空間と記号の空間をいったりきたりすること。この上に、言語によるコミュニケーションや知識の蓄積が構成される。
- ・ いずれも目的は、「いかに少ないサンプルで自由度の高いモデルを同定するか」

身体性を全部すっ飛ばした言語データによるRNN
映像データからの概念理解・知識処理
(Google, Facebook)

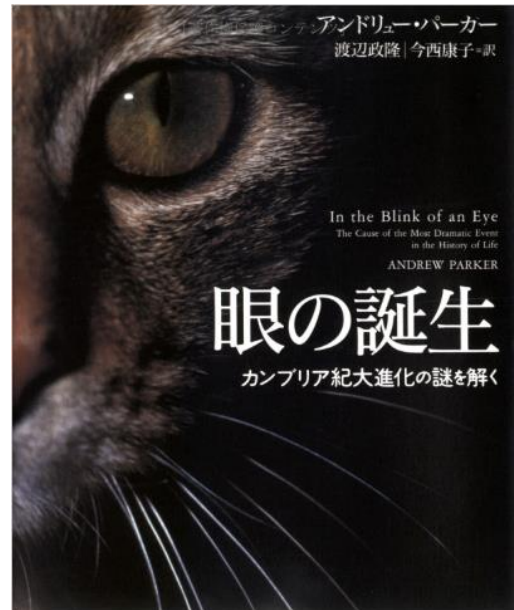


眼の誕生

- カンブリア爆発
 - 5億4200万年前から5億3000万年前の間に突如として今日見られる動物の「門」が出そろった現象
 - 古生物学者アンドリュー・パーカーは、「眼の誕生」がその原因だったの説を提唱
- ディープラーニングにより、見えるようになる
 - さらに、次に何が起こるかを予想して動けるようになる。
- 「眼をもった機械」が誕生する。
 - 機械・ロボットの世界でのカンブリア爆発が起こる。
 - これを日本企業が取れるか？

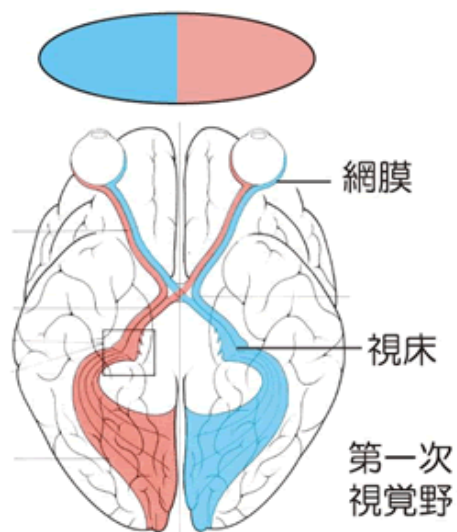


三葉虫: 史上初めて眼をもった生物



53

眼が見える仕組み



←イメージセンサ

←ディープラーニング
(CNN: 畳み込みニューラルネットワーク)

54

既存産業の発展

農業	収穫判定	トラクター、コンバインの適用範囲拡大、効率向上 選別調製等の自動化	自動での収穫 自動での耕うん
建設	測量	掘削、基礎工事、 外装内装作業等の 効率向上	多くの作業の 自動化・効率化
食品 加工	振り分け 確認	カット、皮むき等 の自動化 食洗機に入れる	多くの加工工程の 自動化
組み立て 加工	目視確認の 自動化	動作効率の向上	段取りの自動化 セル生産の自動化
⋮			
A: 画像認識 B: 運動の習熟 C: 計画立案を伴う運動			

55

眼をもった機械・ロボットの典型例 単独の製品から入る

- 農業: トマト収穫ロボット
 - トマトは市場規模も大きく、収穫の工数も大きい。
 - 現状の技術で、トマトの認識ができる。上手にもぎ取ることも可能。
 - 先進的な農場から試しに入れる。
- 建設: 自動溶接機械
 - 建設の工程(例えば溶接)を自動化する
 - 現状の技術で、接合面の状態等の認識ができる。上手に溶接することも可能。機械を当てれば熟練した人でなくとも熟練の人のような溶接ができる。
 - 一部の建設現場で試しに入れる。
- 食品加工: 食洗機にお皿を入れるロボット
 - 食品加工に関わる仕事、まずは食洗機にお皿を入れることを自動化する
 - 現状の技術で、お皿の位置、把持位置の認識ができる。まずは、食器が下げられたところから、食洗機に入れるところを自動化する。(混雑時に重要)
 - ファミリーレストラン等の一部の店舗で試しに入れる。

製品を一刻も早く市場に投入する

56

眼をもった機械・ロボットの典型例 サービス化へ

- 農業: トマト収穫ロボット(異常発見機能つき)
 - トマトの収穫作業に使いながらデータをためる。(ネットワークに接続)
 - 使われる農場数を増やすとともに、コスト削減。
 - 「病気の判定」ができるようになる。それに対して、新たにチャージすることができる。
- 建設: 移動型の自動溶接機械
 - 溶接作業に使いながらデータをためる。(ネットワークに接続)
 - 使われる現場の数を増やすとともに、コスト削減。
 - そのうち、移動しながらの溶接作業や、すでに終わった溶接のチェックができるようになる。それに対して新たにチャージできる。
- 食品加工: お皿管理ロボット
 - お皿を入れながらデータをためる。(ネットワークに接続)
 - 使われる店舗の数を増やすとともに、コスト削減。
 - そのうち、残飯の処理、食器を棚に戻す、食器に盛り付ける等が可能になってくる。それに対して新たにチャージできる。

製品をネットワークにつなぎ、ソフト面の機能向上とともにサービス課金へ

57

眼をもった機械・ロボットの典型例 プラットフォーム化へ

- 農業: トマト栽培管理プラットフォーム
 - トマトの水やり、施肥、害虫駆除等を自動でできるようになる。
 - トマトの収穫全体を管理するプラットフォームを構築できる。このプラットフォームを利用すれば、品質のいいトマトがたくさん取れるというビジネス。
- 建設: 建設現場(躯体工事)プラットフォーム
 - 溶接作業を中心にして、鉄筋を組む作業、コンクリートを入れる作業などを次々に自動化していく。
 - 建設現場の作業全体のプラットフォームを構築できる。プラットフォームを利用すると、自動化された建設現場の施工管理ができるというビジネス
- 食品加工: 調理プラットフォーム
 - 食器の管理から食材管理、調理全般を行うまで自動化していく。
 - 調理全体のプラットフォームを構築できる。プラットフォームを利用すると、店舗の応じた調理が自動で提供されるというビジネス

その「場」の機能全体をプラットフォームとして提供

58

眼をもった機械・ロボットの典型例 海外展開へ

- 農業: グローバルなトマト栽培管理プラットフォーム
 - プラットフォームをそのまま海外に展開し、海外の事業者から課金する。
 - 日本の品質のトマトを、欧米の食卓に届けることができる。
 - 日本はきめ細やかな栽培方法を工夫し続けることで競争力を維持。
- 建設: グローバルな建設現場(躯体工事)プラットフォーム
 - プラットフォームをそのまま海外に展開し、海外の事業者から課金する。
 - 日本のレベルの高い施工技術を、機械化したプラットフォームで提供できる。
 - 日本は建物の耐久性や省エネ、デザイン等、付加価値の高い工夫を続けることで競争力を維持。
- 食品加工: グローバルな調理プラットフォーム
 - プラットフォームをそのまま海外に展開し、海外の事業者から課金する。
 - 日本のレベルの高い食を、世界の飲食店で提供できる。
 - 日本は高い食のレベルを維持し続ける。おいしいものを開発し続ける。

プラットフォームを世界に展開し、高い収益を維持する

59

モノ売りからの脱却へ

- 「眼のある機械」は、データの継続的収集が不可欠
 - 製品からデータが戻るようにしないと、継続的な品質向上につながらない
 - つまり、製品がネットワークに接続されることがほぼ確定している
- すると、眼のある機械の「稼働」に対して課金できるようになる
 - 「学習済みモデル」の品質が上がれば、価格を上げることができる
 - 内部コストを下げれば、利益を上げることができる
 - モノ売りからサービス売りへの転換が容易にできる
- さらに、眼のある機械を起点とする「場」全体のプラットフォーム化へ
 - 製品が置かれるオフィス、家、商業施設、工場、農場、建設現場など、製品が取得するデータ・提供するサービスを起点として、その周りのお金・情報の流れに広げ、事業チャンスをとっていくことができる。
- それを世界展開し、日本品質でサービスを提供する
 - 「学習済みモデル」は日本で作り続け、競争力を維持し続ける

眼のある機械の市場投入→サービス化→周辺を含んだプラットフォーム化
→海外へ大きく展開という流れが王道

60

日本なりのプラットフォーム戦略

- DLの技術はコモディティ化する。
 - 競争力をもつのは、データとハードウェア。
 - 早くDLの技術を取り入れてしまえばよい。
 - DLの技術とハードウェアのすりあわせになった瞬間、日本企業が再度、力を取り戻せる。
- 欧米のスタートアップ(とDL研究者)は、意外なほどハードウェアに対する抵抗感がある
 - そもそも、産業用ロボットの導入台数は日本が(ほぼ)トップ
 - また、ロボットに対する社会的抵抗感もある。
 - 米国は雇用を守らないといけない。日本は人手が足りない。
- ものが関連しないプラットフォームは無理
 - 英語圏でやったほうが絶対に強い。
 - 広告費規模でも10倍、ECの規模でも3倍以上
- 日本には、検索エンジンもECもSNSもあった。
 - GoogleやAmazon, Facebookの位置の企業を出せなくはなかった。
 - しかし、結果はそうになってない。原因は明確で、英語圏でなかったから。

ものづくりを起点に眼をもった機械を作り、
プラットフォーム化するのは日本ならではの戦略

61

機械・ロボットのカンブリア爆発

- 介護施設や病院等での見守り・介護ロボット
- 医療(X線、CT、皮膚、心電図、手術ロボット)
- 警備、防犯技術
- 顔による認証・ログイン・広告技術、表情読み取り技術(サービス業全般に重要)
- 国家の安全保障、入国管理、警察業務、輸出入管理業務における活用
- 防災系(河川、火山、土砂崩れを見張る)
- 重機系(掘削、揚重)、建設現場系(セメント固め、溶接、運搬、取り付け)
- 農業系(収穫、選果、防除、摘花・摘果)
- 自動操縦系(ドローン、小型運搬車、農機、建機)
- 自動運転系、物流
- 産業用ロボット系(特に組み立て加工等)
- 調理系(牛丼、炊飯、ファミリーレストラン、外食全般)
- ペットロボット系
- 片付けロボット(家庭、オフィス、商業施設)
- 新薬発見や新素材の開発(遺伝子の認識・分析、実験ロボット)
- 廃炉系(深海や鉱山、宇宙も含めた極限環境)

農業・建設・食品加工だけでなく、医療や介護、製造、廃炉なども。

62

DLに関わる海外企業

- ベンチャー
 - Deep Mind (英): DLの技術力をもった企業。DQNによるゲーム、アルファ碁、医療など。2011創業。Googleが2014に£ 400Mで買収。
 - Enlitic: 医療画像(X線)におけるDL活用。2014創業、15M調達。
 - Nervana Systems: 医療、農業、金融、自動車、エネルギー等における画像処理。24M調達後、インテルが2016買収。
 - Emotient: 顔の表情を認識する会社。2012創業、6M調達後、Appleが2016買収。
 - Affectiva: 映画やTV番組のどこで表情が変わったのかを読み取る。2009創業、34M調達。
 - Perceptio: DLによる写真分類アプリ開発。創業、調達額不明。Appleが2015買収。
 - VocalIQ (英): DLによる音声認識。1M調達後、Appleが2015買収。
 - Atomwise: ドラッグディスカバリーへのDL活用。新薬の候補物質を見つける。YC卒業生。2012創業。6M調達。
 - Descartes Labs: DLによる衛星画像の分析。農業への適用。2014創業、8M調達。
 - Canary: DLによるホームセキュリティ。2012創業。41M調達。
 - Netatmo: 家電。DLによる監視カメラも。2011創業。38M調達。
 - Pilot AI Labs: DLの画像認識を使ったドローンの自動操縦。まだ小さいが、動画が面白い。
 - MetaMind: 画像認識一般。2014創業、8M調達。
 - SkyMind: JavaベースのDL提供。2014創業。3M調達。
 - AlchemyAPI: DLによる言語処理と画像認識。クラウドで提供。2005創業、2M調達。IBMが2015に買収。
- ZenRobotics (フィンランド): ごみの選別ロボット。2007創業、17M調達。
- 製造業
 - GE: DLによる医療画像の診断。
 - Dyson: 掃除機に眼をつけたものを開発。インペリアル・カレッジにRoboticsラボ設立。
 - Kuka (独): 産業用ロボットへのDL適用
 - Mobileye (蘭): 車用の画像認識を提供。1999イスラエルで創業。2014年上場。時価総額10B。
 - LG (韓): インチョン空港で、DLを使ったロボットでの案内の実験
 - Samsung (韓): DLを使った胸部エコー検診の医療機器

急がないといけない

63

どこから始めるか

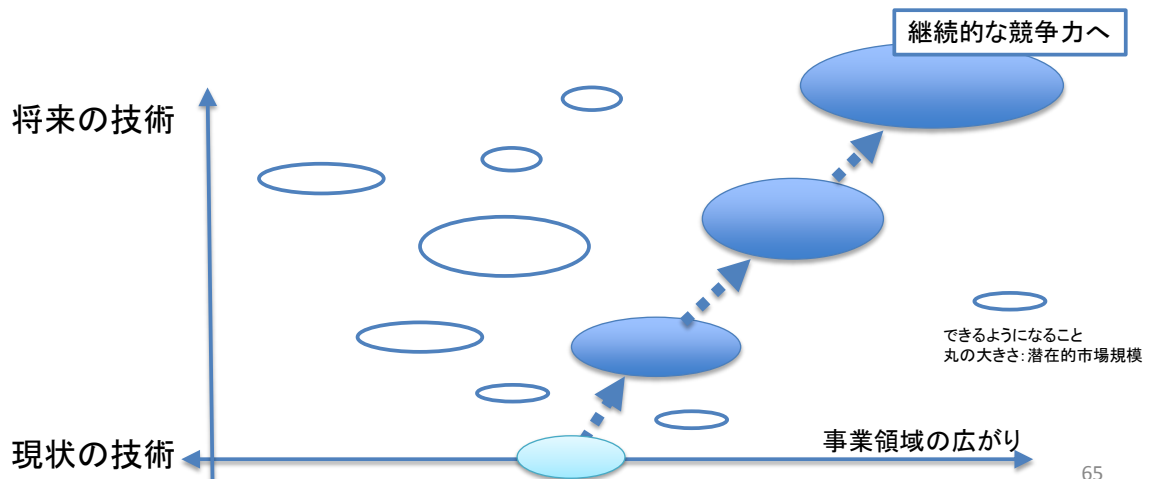
- 始めること自体はそれほど難しくない
 - 数学の知識: 線形代数や最適化
 - プログラムの知識: python, GPU
- ライブラリが揃っている
 - Tensorflow: Google, python
 - Caffe: UCバークレー, C++ベース
 - ほかにも, Keras, chainer, torch7, ...
- 教科書
 - 入り口: 「人工知能は人間を超えるか」(拙著)
 - 読み物: 「人間さまお断りー人工知能時代の富と仕事の手引き」(Jerry Kaplan, 近刊)
 - 教科書: 「深層学習」(岡谷貴之)
 - 教科書: 「Deep Learning」(Y. Bengioら, MIT pressから来年出版。翻訳を出版予定)
 - www.deeplearningbook.org にPDFがあり、読めます。
 - あとは論文を読んでください。
- 3つの主要な国際会議
 - ICML, NIPS, ICLR (それぞれ年1回)
 - ウェブで全ての情報が見れます。

理系の人が半年もやればそこそこできるようになる

64

経営的な側面からみた考え方

- 売上が5%, 10%増えるかも知れなく、(大企業であっても)5倍, 10倍になるかもしれないことを考える
 - 「認識」や「運動の習熟」で何ができるようになるか
 - いままでの工業化は機械が認識できない大前提で、「環境を整え」自動化。**その大前提を外して考える**が必要
- ニーズ駆動のイノベーションと混同しない
 - ここ10-20年の情報系はニーズ探しだった。Google, Facebook, Amazon, Uber, Airbnb, ...
 - 今回は、圧倒的な性能向上。**計画して投資する。**



65

ポイント

- 製品を作って、その利益を再投資に回すサイクルをいかに作れるか。
 - 最終的な競争力は技術にはない。**データとハードウェア。**
 - インターネットの世界で起こったことと同じ。
 - 個々の製品の性能向上→サービス化・プラットフォーム化
- 最も強いインセンティブをもつ企業が、最も効率的に再投資のループを作り勝つ
 - 例えば、生産の効率化ができるとしても、そこに対してもっと強いインセンティブを持つ企業が他にいないのか？
 - 顔認識ができるとして、それに対して最も強いインセンティブをもつ企業はどこなのか？
 - つまり、自社事業と**「最も直接的」に関係する技術革新**を考えないといけない。
- 実はこれは社内文化との戦い
 - 情報系はこれまでも「本流」から外れ、弱かった
 - そのなかでも「機械学習」の技術をもつ部署をいかに中心にもってこれるか。長年の伝統だった「本流」をいかに破壊し再構成するか。
 - それを**トップダウン**で**意思決定**できるか。

自社が最も強いインセンティブを感じる製品・サービスを描き、それをトップダウンに実行できるか

66

必要になる「学習工場」

- 学習工場のイメージ
 - 機械学習を使える高度な人材
 - 高性能な計算機
 - データを準備する環境
- 学習工場で出荷されるもの
 - 学習済みの「モデル」が作られる。頭脳の部分。
 - これが最終製品に載せて出荷される。
 - それぞれの事業者が学習工場をもつ。
- 「工場」なので、数十億～数千億円(数兆円)の投資規模
 - R&Dではない。最終製品を作るものである。
 - 10倍生産したいのであれば10倍の規模に。
 - この規模でようやく諸外国と勝算ある状態で戦える
- これを新たな設備投資の形とすることができないか。

67

それぞれに必要なもの

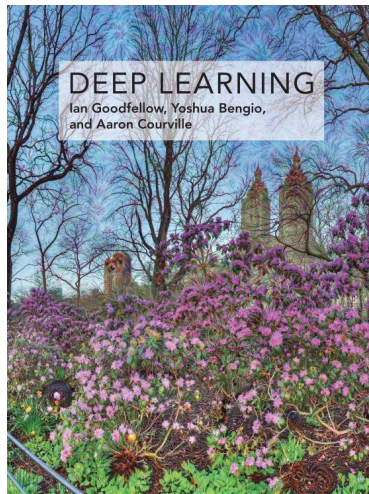
- 学習用のデータづくり
 - アノテーション付きのデータをたくさん作り出す事業者(サプライヤー)
 - そこには大量の雇用が発生
 - 2次サプライヤー
 - そのためのツールを開発する事業者
 - そのためにクラウドソーシングを活用する事業者
 - 模擬データを出すためのシミュレータを開発する事業者
 - 極限環境でのデータをとる事業者
- 計算機
 - DL用の計算機: 海外が強い(NVIDIA)
 - ライブラリ: 海外が強い。TensorflowやKerasなど。
 - アルゴリズム: 海外が強い。GoogleやFacebookなど。
 - ひとまずしっかりと輸入。あまり国産にこだわっても良くない。
- 学習職人(高度なスキル・知識をもった人材)
 - いま、これが圧倒的に足りない。
 - 大学で人材育成を急ぐ。企業内での人材研修。
 - 研修プログラムの提供、資格制度作りを急ぐべき。

大量の雇用、
たくさんの事業者を生む
産業としての広がりは大きい

しっかり輸入
(いずれ独自の競争力)

ハイスpekク人材

68



Ian Goodfellow, Yoshua Bengio,
and Aaron Courville, MIT press 2016.11

- 人工ニューラルネットワークの最初の実験が1950年代に行われたのに、なぜ最近になってようやく、深層学習が極めて重要な技術と認識されるようになったかは、不思議に思うかもしれない。深層学習は、1990年代から商業的な応用ではうまく使われていたが、最近まで、しばしば技術というよりはアートであり、専門家だけが使えるものと見なされた。深層学習のアルゴリズムでよい性能を得るには、必要なスキルがあることは真実である。幸いにも、必要なスキルの量は、訓練データの量が増えるにつれて減っている。今日、複雑なタスクで人間の性能に到達する学習アルゴリズムは、1980年代におもちゃの問題を解くのに苦労した学習アルゴリズムとほとんど同一である。これらのアルゴリズムで訓練するモデルは、とても深いアーキテクチャでの訓練を簡略化する変化をしてはいるが。最も重要な新しい進歩は、今日ではアルゴリズムが成功するのに必要とするだけのリソースを、アルゴリズムに提供することができることである。

- 2016年の時点で、大体の目安として、教師ありの深層学習のアルゴリズムは、一般的に、カテゴリごとに約5000のラベル付き事例で、許容できる性能を達成し、少なくとも1000万のラベル付きの事例を含むデータセットで訓練すれば、人間の性能と匹敵する、あるいは超える。これよりも小さいデータセットでうまくいくことは、重要な研究分野であり、教師なしあるいは半教師あり学習で、大量の教師なし事例をうまく活かす方法に特に焦点が当てられる。

69

今回の人工知能ブームは2階建て： 可能性と限界をしっかりと理解すべき

- 多くの場合、「デジタル革命」を指している。
- アルファ碁や画像認識など、「ディープラーニング革命」の出来事とまざって理解されている。
- 長期的に重要なこと：デジタル革命。ここ20年のGDPの伸びの大半はIT技術由来。
- 戦略的チャンス：ディープラーニング革命。特に「眼のある機械」は、製造業の切り札になり得る。

ディープラーニング
革命

「眼の誕生」：
まさにいまこの瞬間に起こっているイノベーション

デジタル革命
(人工知能という名の情報技術の擬人化)

米国から20年遅れのイノベーション
ようやく「情報技術」「データ化」の大きな可能性に
人工知能という「擬人化」によって、人々が気づき始めた。
重要であることは論をまたないが、それだけでもはや競争
力にはならない。

変わりゆく社会

- 倫理や社会制度の議論がもう一度必要になる
 - 自動運転で危険回避のときは？人の命の重さは？
- 人工知能システムが社会に広がったときの不具合の問題
 - 製造者責任？保険や社会保障のほうが適切では
- 心をもつように見える人工知能を作ってよいか
 - プログラムの停止させると悲しむ？
 - 恋愛させるビジネスなど(映画「Her」の世界)
- 人工知能を使った軍事
 - ロボット兵士やドローン
 - 権力者を倒す、心を操る？
- 人工知能が知財を生み出す場合の権利
 - 著作権や特許は認めるべきか
- 実は人間が本来的にもっている権利がもっとあるのではないか
 - 忘れられる権利、見られない権利、大目に見られる(警告を受ける)権利、好きになる権利、...



人工知能学会 倫理委員会
(松尾が委員長)でも議論。
社会全体で議論していく必要。
内閣府でも、「人工知能と人間社会に
関する懇談会」にて議論。

71

人工知能学会 倫理指針

2017.2.28

序文

人工知能研究は、人間のような知性を持ち自律的に学習し行動する人工知能の実現を目指している。人工知能が、産業、医療、教育、文化、経済、政治、行政など幅広い領域で人間社会に深く浸透することで、人々の生活が格段に豊かになることが期待される一方で、悪用や濫用で公共の利益を損なう可能性も否定できない。

高度な専門的職業に従事する者として、人工知能の研究、設計、開発、運用、教育に広く携わる人工知能研究者は、人工知能が人間社会にとって有益なものとなるようにするために最大限の努力をし、自らの良心と良識に従って倫理的に行動しなければならない。人工知能研究者は、社会の様々な声に耳を傾け、社会から謙虚に学ばなければならない。人工知能研究者は技術の進化及び社会の変化に伴い、人工知能研究者自身の倫理観を発展させ深めることについて不断の努力をおこなう。

人工知能学会は、自らの社会における責任を自覚し、社会と対話するために、人工知能学会会員の倫理的な価値判断の基礎となる倫理指針をここに定める。学会員はこれを指針として行動するよう心がける。

- 1 (人類への貢献) 人工知能学会会員は、人類の平和、安全、福祉、公共の利益に貢献し、基本的人権と尊厳を守り、文化の多様性を尊重する。人工知能学会会員は人工知能を設計、開発、運用する際には専門家として人類の安全への脅威を排除するように努める。
- 2 (法規制の遵守) 人工知能学会会員は専門家として、研究開発に関わる法規制、知的財産、他者との契約や合意を尊重しなければならない。人工知能学会会員は他者の情報や財産の侵害や損失といった危害を加えてはならず、直接的のみならず間接的にも他者に危害を加えるような意図をもって人工知能を利用しない。
- 3 (他者のプライバシーの尊重) 人工知能学会会員は、人工知能の利用および開発において、他者のプライバシーを尊重し、関連する法規に則って個人情報の適正な取扱いを行う義務を負う。
-
- 8 (社会との対話と自己研鑽) 人工知能学会会員は、人工知能に関する社会的な理解が深まるよう努める。人工知能学会会員は、社会には様々な声があることを理解し、社会から真摯に学び、理解を深め、社会との不断の対話を通じて専門家として人間社会の平和と幸福に貢献することとする。人工知能学会会員は高度な専門家として絶え間ない自己研鑽に努め自己の能力の向上を行うと同時にそれを望む者を支援することとする。
- 9 (人工知能への倫理遵守の要請) 人工知能が社会の構成員またはそれに準じるものとなるためには、上に定めた人工知能学会員と同等に倫理指針を遵守できなければならない。

73

個人にとって: 職業の変化

- 10年～20年で、日本の労働人口の49%が人工知能やロボット等で代替可能に(NRI調べ、2015年12月)
- 新しい仕事が出てくる。

人工知能やロボット等による代替可能性が高い100種の職業 (50音順、並びは代替可能性確率とは無関係)

※職業名は、労働政策研究・研修機構「職務構造に関する研究」に対応

IC生産オペレーター	こん包工	電子計算機保守員 (IT保守員)
一般事務員	サッシ工	電子部品製造工
鋳物工	産業廃棄物収集運搬作業員	電車運転士
医療事務員	紙器製造工	道路バトロール隊員
受付係	自動車組立工	日用品修理ショップ店員
A V・通信機器組立・修理工	自動車塗装工	バイク便配達員
駅係員	出荷・発送係員	発電員
N C研削盤工	じんがい収集作業員	非破壊検査員
N C旋盤工	人事係事務員	ビル施設管理技術者
会計監査係員	新聞配達員	ビル清掃員
加工紙製造工	診療情報管理士	物品購買事務員
貸付係事務員	水産ねり製品製造工	プラスチック製品成形工
学校事務員	スーパー店員	プロセス製版オペレーター
カメラ組立工	生産現場事務員	ボイラーオペレーター
機械木工	製パン工	貿易事務員
寄宿舎・寮・マンション管理人	製粉工	包装作業員
CADオペレーター	製本作業員	保管・管理係員
給食調理人	清涼飲料ルートセールス員	保険事務員
教育・研修事務員	石油精製オペレーター	ホテル客室係
行政事務員 (国)	セメント生産オペレーター	マシンニングセンター・オペレーター
行政事務員 (県市町村)	繊維製品検査工	ミシン縫製工
銀行窓口係	倉庫作業員	めっき工
金属加工・金属製品検査工	惣菜製造工	めん類製造工
金属研磨工	測量士	郵便外務員
金属材料製造検査工	宝くじ販売員	郵便事務員
金属熱処理工	タクシー運転者	有料道路料金収受員
金属プレス工	宅配便配達員	レジ係
クリーニング取次店員	鍛造工	列車清掃員
計器組立工	駐車場管理員	レンタカー営業所員
警備員	通関士	路線バス運転者
経理事務員	通信販売受付事務員	
検収・検品係員	積卸作業員	
検計員	データ入力係	
建設作業員	電気通信技術者	
ゴム製品成形工 (タイヤ成形を除く)	電算写植オペレーター	

人工知能やロボット等による代替可能性が低い100種の職業 (50音順、並びは代替可能性確率とは無関係)

※職業名は、労働政策研究・研修機構「職務構造に関する研究」に対応

アートディレクター	児童厚生員	バーテンダー
アウトドアインストラクター	シナリオライター	俳優
アナウンサー	社会学研究者	はり師・きゅう師
アロマセラピスト	社会教育主事	美容師
犬訓練士	社会福祉施設介護職員	評論家
医療ソーシャルワーカー	社会福祉施設指導員	ファッションデザイナー
インテリアコーディネーター	獣医師	フードコーディネーター
インテリアデザイナー	柔道整復師	舞台演出家
映画カメラマン	ジュエリーデザイナー	舞台美術家
映画監督	小学校教員	フラワーデザイナー
エコノミスト	商業カメラマン	フリーライター
音楽教室講師	小児科医	プロデュサー
学芸員	商品開発部員	ベンション経営者
学校カウンセラー	助産師	保育士
観光バスガイド	心理学研究者	放送記者
教育カウンセラー	人類学者	放送ディレクター
クラシック演奏家	スタイリスト	報道カメラマン
グラフィックデザイナー	スポーツインストラクター	報道カメラマン
ケアマネージャー	スポーツライター	法務官
経営コンサルティング	声楽家	マーケティング・リサーチャー
芸能マネージャー	精神科医	ミュージシャン
ゲームクリエイター	ソムリエ	メイクアップアーティスト
外科医	大学・短大教員	盲・ろう・養護学校教員
言語聴覚士	中学校教員	幼稚園教員
工業デザイナー	中小企業診断士	理学療法士
広告ディレクター	ツアーコンダクター	幼歯科教員
国際協力専門家	ディスクジョッキー	料理研究家
コピーライター	ディスプレイデザイナー	旅行会社カウンター係
作業療法士	デスク	レコードプロデューサー
作家	テレビカメラマン	レストラン支配人
作曲家	テレビタレント	録音エンジニア
雑誌編集者		
産業カウンセラー		
産婦人科医		
歯科医師		
	ネイル・アーティスト	

重要になる人間の役割

- 対人間のコミュニケーション
 - 低付加価値のサービスは機械化・ロボット化
 - 高付加価値のサービスは人間が
 - コミュニケーション力や人間力、教養が重要に(時代を超えて重要)
- 目的の設定・価値判断・責任主体
 - なにが大事か。異なる価値観のものをどうバランスさせるのか。
 - ニーズを捉える企画や経営
 - 「謝る」ことは代替できるのか
- センサー・創造性
 - 「生命」由来の、ものまねではない創造性
 - テスター: たぶん社会保障としての仕事はこれになる
- 人工知能・ロボットを使う仕事
 - 人工知能・ロボットの開発・運用
 - 開発側に回るか、ユーザ側に回るかは極めて重要(まさにいまの検索やSNSと同じ)
 - 数理能力+プログラミング能力+ハードウェア構成能力
 - 教える(教師データを与える)、補完する、伝える
 - 熟練の技術、知識、スキル

75

我々はどういう社会を作りたいのか

- 人工知能技術が進めば進むほど、「与えられた目的」に対して、それを実現する手段は賢くできるようになる。
 - 人間＝知能＋生命
 - 知能は、目的を与えられたときの問題解決の力。
 - 生命は、目的を持つ。
 - 自己保存、自己複製、仲間を守るなど。そうしないものは、進化の過程で滅んできたため。
 - 人工知能の技術は知能、すなわち問題解決の技術。
- そうすると、与える目的自体の是非の議論のほうがより重要になる。
 - なにが社会で大事なのか？
 - 個人の幸せや社会全体の幸せはどのように考えればいいのか？
 - 異なる価値観のものをどのようにバランスさせればいいのか？
- これまで人文社会学系でされてきたような議論が今後、改めて重要になる。
 - 特に、哲学、政治学、社会学、法学、心理学、経済学など。

我々は、どういう社会を作りたいのか？
幸せで持続可能な社会とは？

76

日本の戦略

日本の社会課題に対して、DLとものづくりの掛けあわせによる「眼をもった機械」を開発し解決する

- 農業分野に「眼をもった機械」を適用することで
 - 休耕地が耕せる。除草・防除や収穫ができる。収量が増える。
- 介護分野に適用することで
 - 介助も楽に。移動したりトイレにいけるようになり、より自立した生活ができる。
- 廃炉作業に適用することで
 - 危険な状況で人が作業しなくてよくなる。工期を短縮できる。
- 河川や火山を見張ることで
 - 河川の氾濫や土砂崩れ、噴火などの危険な状態・予兆を早期に発見できる。
- こうした技術を使った製品を海外に展開していくことで
 - 新たな輸出産業に。GDPの増加につながる。

地方からグローバルへ

労働の必要な地方を舞台に技術を伸ばす。それをグローバルに展開

77

ディープラーニング×ものづくり： 「眼をもった機械」による日本の新たな産業競争力の実現へ

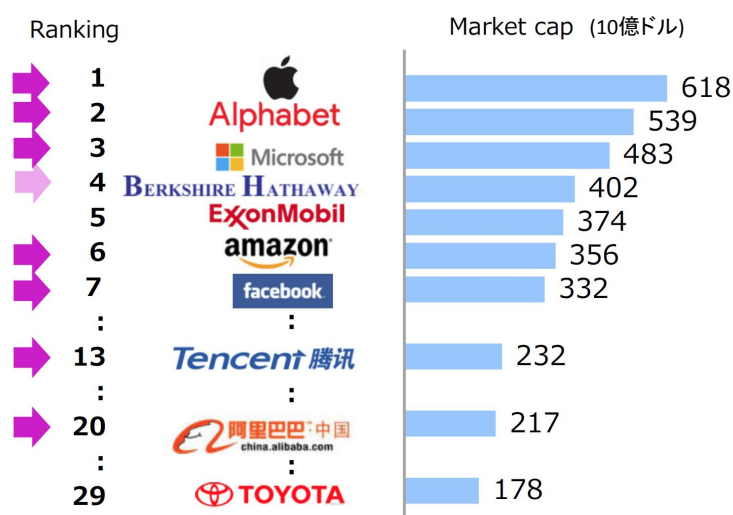
- 少子高齢化しており、労働力が不足している。
 - 頭脳労働は不足していない。「運動を伴う労働」のニーズが高い。
 - 農業従事者、建設・物流、介護、廃炉、熟練工の後継者、etc
- ディープラーニングが解決策になり得る。
 - 眼をもった機械：認識や運動の上達ができる機械・ロボット
 - ものづくりと相性がよく、日本の強みを活かせる。素材や駆動系も強い
 - 眼をもった機械というのは世界的にはもう常識。しかも自動運転だけではない。
- 新たな投資概念の必要性
 - 人への投資をいかに既存の枠組み・文化を踏まえてやるか
 - 企業からの投資を引き出し、そこに人材が流入する生態系を作るか
 - 例えば「学習工場」
- チャンスを捉えるには、正しく早く動いていくことが重要
 - ディープラーニング人材の育成
 - 事業・産業がどう変わるかを早期に検討
 - 社会全体で新しい未来像を描いていくこと

78

APPENDIX

79

時価総額ランキング



Google (Alphabet) 研究開発費 1兆4000億円
 ここ20年、日本はITでは負け続け。もう「投了寸前」。
 眼のある機械(ディープラーニング×ものづくり)は、最後のチャンスではないか。

「逆算」で考えるプレイヤーが日本にはほとんどいない

- 眼のある機械でどれだけの新しい市場が広がっているのか？
- そこから逆算すれば、どのくらい投資してもいいのか？
- DLの人材が市場でどのくらいで評価されているのか
- そこから逆算すれば、どのくらいコストをかけてもいいのか？

81

3つの変化

- 1. 純粋な技術変化
 - ディープラーニング:「認識」ができるようになった
 - トランジスタ:電気を増幅できるようになった
- 2. 技術環境の変化
 - 通信速度があがった
 - ストレージの値段が下がった
 - スマホを全員もつようになった
- 3. 社会受容の変化
 - 自分の写真をアップしても平気になった
 - ネットで買い物しても不安が平気になった
 - 検索上位に来るのが重要になった



めったにない
影響を考えるのは簡単
まじめに事業計画を書く

ここ10-20年は、2とか3の話が中心だった。ニーズ主導のイノベーションだった日本人は苦手だった。
今回は1. まじめに考えれば分かる。

82

やるべきこと

- 「眼のある機械」という新たなカテゴリの機械・ロボットによる潜在市場をきちんと調査する。
- どの市場を取るべきかを明確にし、事業戦略を立てる。それに向けての技術投資をする。
- DLの技術をもった若手を増やす。
 - 逆算で給料を設定すればあり得ないくらいの好条件にできるはず
- そうした若手を中心に事業戦略上、必要な技術開発を超スピードでやる。

この技術の中心は20代(30代)。DLができる人材をとにかく増やす。
上の人にはうるさいことを言わない。(上への説明コストがめっちゃスピードを削ぐ。)
ただし、潜在市場を見極める作業と投資の意思決定は全社でしっかりやるべき。

83

健康医療分野における活用の可能性

- 画像診断系(遠隔医療含む)
 - レントゲン/CT/MRI
 - 内視鏡
 - 眼底、皮膚病、病理診断
- 薬の管理系
 - 調剤薬局
 - 病院内の薬の配送・管理
 - 画像認識による常用薬の同定と飲み合わせチェック
 - 病院内で正しく飲んだかのチェック
- 見守り系
 - 病院内、家庭内での認知症の見守り
 - ナースコールの誤アラーム防止
- 業務効率系
 - 音声認識による看護師／医師の業務時間短縮(看護記録、電カルへの入力等)
 - 内科の問診等のテンプレート化

84

参考：画像診断系（ニュース等より）

- ディープマインド
 - ロンドンに本社を置いているが、ロンドンのムアフィールズ眼科病院と提携し、100万枚以上の網膜画像を学習した。眼底検査は、血管や神経を外から観察できる唯一の器官であり、画像認識との相性が非常に良い。糖尿病性網膜症と加齢黄斑変性を診断するアルゴリズムは、すでに専門家と同程度か、それ上回る判定を達成している。
- IBM
 - 眼底に対する医療画像で同様の取り組みを行っている。糖尿病性網膜症に関して、重症度分類（5段階）の精度が87%に達したと発表した。網膜画像を取りさえすれば、かなり高い精度で症状が判断できるということである。
- 米国エンリティック社
 - レントゲン写真やCTスキャンから悪性腫瘍を見つけ出す。同社が開発した悪性腫瘍の検出システムは、放射線医師を上回る精度だという。肺がんに関する画像データベース「LIDC(Lung Image Database Consortium)」や「NLST(National Lung Screening Trial)」を用いて検証したところ、同社のシステムの肺がんの検出の精度は、放射線医師が1人だけで肺がんを検出する精度を5割以上も上回った。

85

参考：画像診断系（ニュース等より）

- GEヘルスケア
 - カリフォルニア大学サンフランシスコ校と、ディープラーニングを使ったライブラリに関する提携を結んだ。診断の結果をGEのイメージング機器で展開しようとしている。
- スタンフォード大学のグループ
 - 皮膚がんの特定に使った。ネイチャーに出した論文によると、この手法が経験を積んだ皮膚科医と同じように皮膚がんを特定できるとのことである。
- アルタリス（Arterys、サンフランシスコにある株式非公開の医用画像会社）
 - 開発したソフトウェアは、すでにFDA（米国食品医薬品局）によって販売が許可された。アルタリスのアルゴリズム「DeepVentricle」は心室内の輪郭を写したMRI画像を分析し、患者の心臓が保持・拍出できる血液量を計算する。
- 韓国天安公衆衛生センター
 - 脳の画像からアルツハイマーを予測する研究が行われている。精度81%。
- エルピクセル
 - 脳動脈瘤をMRIからを見つけることを行っている。病理医の診断をサポートする。

86

個人情報保護法の概要

平成 2 9 年 9 月 2 6 日
個人情報保護委員会事務局

目次

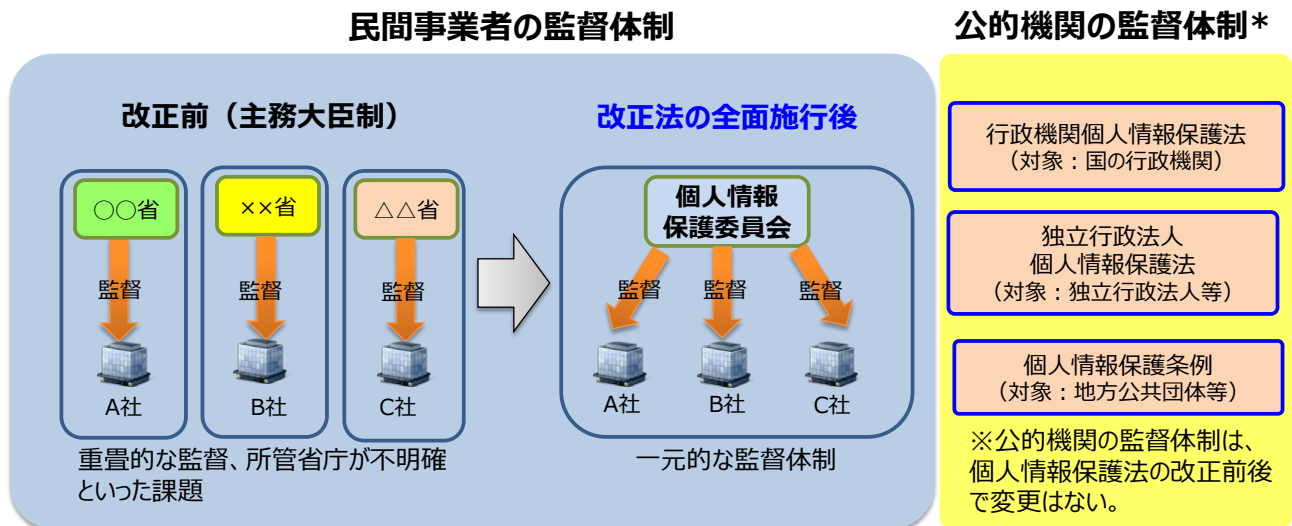
はじめに

1. 個人情報保護法とは
2. 個人情報とは
3. 事業者が守るべきルール
4. 匿名加工情報とは
5. お役立ち情報

はじめに

○個人情報保護委員会とは ※青字箇所は改正部分を示しています。

- 内閣府の外局、高い独立性のある行政機関
- 主務大臣が有していた監督権限を個人情報保護委員会へ一元化
- 事業者に対して、必要に応じて報告を求めたり立入検査を行うことができる
また、実態に応じて、指導・助言、勧告・命令を行うことができる



1

はじめに

○個人情報保護委員会とは

沿革

- 平成26年1月1日 特定個人情報保護委員会 設置
(行政手続における特定の個人を識別するための番号の利用等に関する法律第36条)
- 平成28年1月1日 個人情報保護委員会 設置
(特定個人情報保護委員会から改組)
(個人情報の保護に関する法律及び行政手続における特定の個人を識別するための番号の利用等に関する法律の一部を改正する法律第1条及び第4条)

所掌事務

(第61条)

- ①マイナンバー制度に関する事務（監視・監督、特定個人情報保護評価）
- ②個人情報保護法に関する事務、監視・監督（個人情報保護法を所管）
- ③上記①、②に共通する事務（広報・啓発、国際協力等）

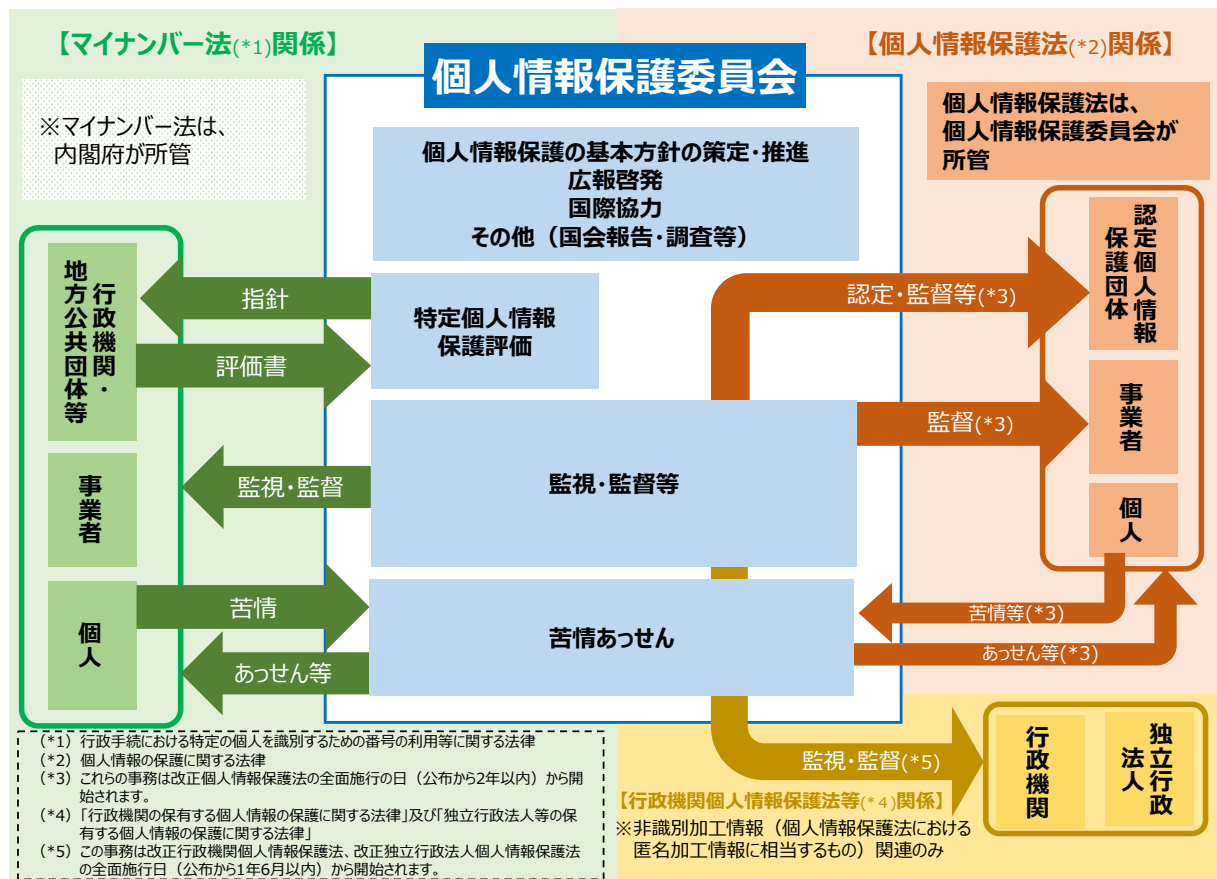
組織

- 委員長1名・委員8名（合計9名）の合議制（行政委員会）
- 委員長・委員は独立して職権を行使（任期5年）
- 委員会事務局の職員数：125名（平成29年7月現在）



2

はじめに



3

はじめに

(参考) 個人情報保護法ガイドライン

- 改正個人情報保護法の全面施行に伴い、同法の監督権限が当委員会に一元化されたことから、当委員会が、全ての分野に共通に適用される汎用的なガイドラインを定める(平成28年11月30日公布)
「通則編」(個人情報保護法全体の解釈・事例)、「匿名加工情報編」、
「個人データの第三者提供時における確認・記録義務編」、
「外国にある第三者への個人データの提供編」
- 一部の分野については、個人情報の性質及び利用方法並びに現行の規律の特殊性等を踏まえて、上記のガイドラインを基礎として、当該分野において更に必要となる別途の規律を定める
「金融関連(信用等含む)」
⇒金融分野、信用分野又は債権管理回収業分野において、さらに必要となる別途の規律を定めたもの
※それぞれ、金融庁、経済産業省及び法務省と個人情報保護委員会の連名のガイドライン
「医療関連」
⇒現場の実態に配慮した規律や事例を記載したもの
「情報通信関連」
⇒通信の秘密・信書の秘密等、業法上必要な規律を含む別途の規律を定めたもの
※総務省のガイドライン

4

1.個人情報保護法とは

- 個人の権利・利益の保護と個人情報の有用性とのバランスを図るための法律
- 基本理念を定めるほか、民間事業者の個人情報の取扱いについて規定



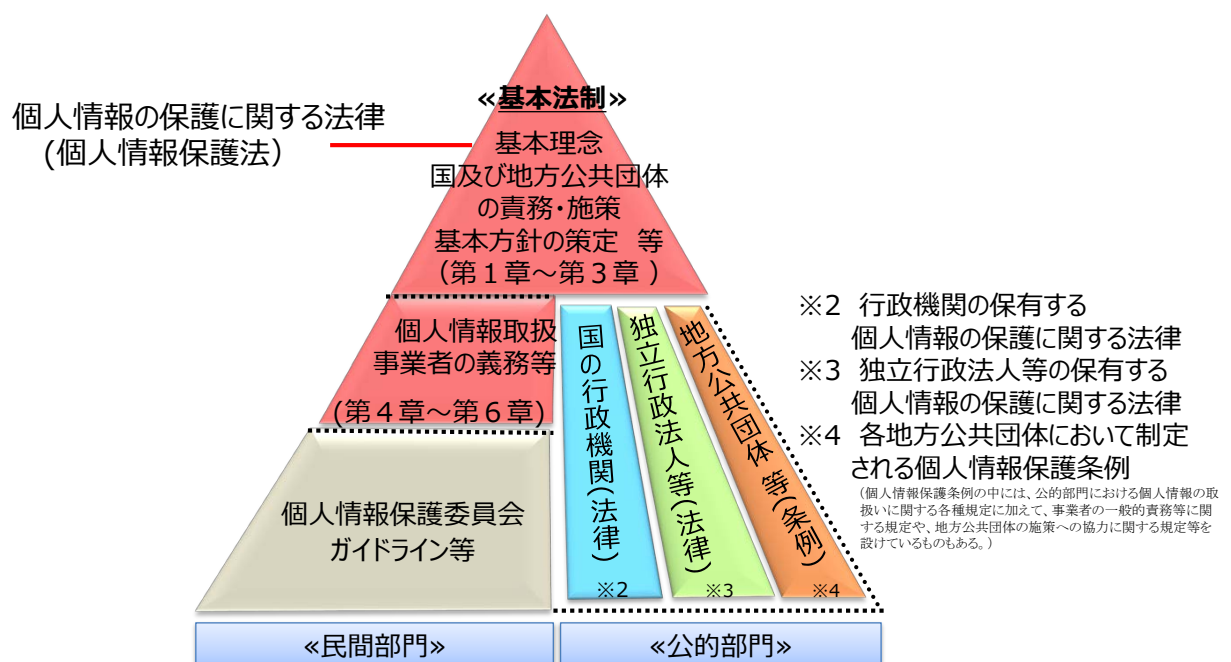
個人情報保護法の目的

第1条

この法律は、高度情報通信社会の進展に伴い個人情報の利用が著しく拡大している ことに鑑み、個人情報の適正な取扱いに関し、基本理念及び政府による基本方針の作成その他の個人情報の保護に関する施策の基本となる事項を定め、国及び地方公共団体の責務等を明らかにするとともに、個人情報を取り扱う**事業者の遵守すべき義務等を定める**ことにより、個人情報の適正かつ効果的な活用が新たな産業の創出並びに活力ある経済社会及び豊かな国民生活の実現に資するものであることその他の**個人情報の有用性に配慮しつつ、個人の権利利益を保護することを目的とする。**

5

1.個人情報保護法とは



6

1.個人情報保護法とは

2003年「個人情報の保護に関する法律」成立（2005年全面施行）

環境の変化

情報通信技術の発展により、制定当時には想定されなかったパーソナルデータの利活用が可能に

1. グレーゾーンの拡大

個人情報に該当するかどうかの判断が困難ないわゆる「グレーゾーン」が拡大

2. ビッグデータへの対応

パーソナルデータを含むビッグデータの適正な利活用ができる環境の整備が必要

3. グローバル化

事業活動がグローバル化し、国境を越えて多くのデータが流通

7

1.個人情報保護法とは

○平成27年9月 改正個人情報保護法が成立（施行は平成29年5月30日）

●改正のポイント●

1. 個人情報保護委員会の新設

個人情報取扱事業者に対する監督権限を各分野の主務大臣から委員会に一元化。

2. 個人情報の定義の明確化

- ①利活用に資するグレーゾーン解消のため、個人情報の定義に身体的特徴等が対象となることを明確化。
- ②要配慮個人情報（本人の人種、信条、病歴など本人に対する不当な差別又は偏見が生じる可能性のある個人情報）の取得については、原則として本人同意を得ることを義務化。

3. 個人情報の有用性を確保（利活用）するための整備

匿名加工情報（特定の個人を識別することができないように個人情報を加工した情報）の利活用の規定を新設。

4. いわゆる名簿屋対策

- ①個人データの第三者提供に係る確認記録作成等を義務化。（第三者から個人データの提供を受ける際、提供者の氏名、個人データの取得経緯を確認した上、その内容の記録を作成し、一定期間保存することを義務付け、第三者に個人データを提供した際も、提供年月日や提供先の氏名等の記録を作成・保存することを義務付ける。）
- ②個人情報データベース等を不正な利益を図る目的で第三者に提供し、又は盗用する行為を「個人情報データベース提供罪」として処罰の対象とする。

5. その他

- ①取り扱う個人情報の数が5000以下である事業者を規制の対象外とする制度を廃止。
- ②オプトアウト（※）規定を利用する個人情報取扱事業者は所要事項を委員会に届け出ることを義務化し、委員会はその内容を公表。（※本人の求めに応じて当該本人が識別される個人データの第三者への提供を停止する場合、本人の同意を得ることなく第三者に個人データを提供することができる。）
- ③外国にある第三者への個人データの提供の制限、個人情報保護法の国外適用、個人情報保護委員会による外国執行当局への情報提供に係る規定を新設。

8

2.個人情報とは

○個人情報の定義

【改正前】

第2条 この法律において「個人情報」とは、生存する個人に関する情報であつて、当該情報に含まれる氏名、生年月日その他の記述等により特定の個人を識別することができるもの（他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるものを含む。）をいう。

【改正後】

第2条 この法律において「個人情報」とは、生存する個人に関する情報であつて、次の各号のいずれかに該当するものをいう。

- 一 当該情報に含まれる氏名、生年月日その他の記述等（文書、図画若しくは電磁的記録（電磁的方式（電子的方式、磁気的方式その他の知覚によっては認識することができない方式をいう。次項第二号において同じ。）で作られる記録をいう。第十八条第二項において同じ。）に記載され、若しくは記録され、又は音声、動作その他の方法を用いて表された一切の事項（個人識別符号を除く。）をいう。以下同じ。）により特定の個人を識別することができるもの（他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるものを含む。）

- 二 個人識別符号が含まれるもの

※青字：改正部分 下線：変更のない部分

9

2.個人情報とは

○個人識別符号とは

- 個人情報の定義の明確化を図るため、その情報単体でも個人情報に該当することとした「**個人識別符号**」の定義を設けた。（第2条2項）
- 「個人識別符号」は以下①②のいずれかに該当するものであり、政令・規則で個別に指定されている。（政令第1条、規則第3、4条）
 - ① 身体の一部の特徴を電子計算機のために変換した符号
⇒DNA、顔、虹彩、声紋、歩行の態様、手指の静脈、指紋・掌紋
 - ② サービス利用や書類において対象者ごとに割り振られる符号
⇒公的な番号
旅券番号、基礎年金番号、免許証番号、住民票コード、マイナンバー、各種保険証等



10

3.事業者が守るべきルール

① 個人情報取得・利用する時のルール

⇒個人情報を取得した場合は、その利用目的を本人に通知、又は公表すること（あらかじめ利用目的を公表している場合を除く。）

② 個人情報を保管する時のルール

⇒情報の漏えい等が生じないように安全に管理すること

③ 個人情報を他人に渡す時のルール

⇒個人情報を本人以外の第三者に渡すときは、原則として、あらかじめ本人の同意を得ること

④ 個人情報を外国にいる第三者に渡す時のルール

⑤ 本人から個人情報の開示を求められた時のルール

⇒本人からの請求に応じて、個人情報を開示、訂正、利用停止等すること

11

3.事業者が守るべきルール

（参考）本人に「通知」、「公表」、「同意」とは

➤ 本人に「通知」とは

⇒本人に直接知らしめることをいい、事業の性質及び個人情報の取扱状況に応じ、内容が本人に認識される合理的かつ適切な方法によらなければならない。

➤ 本人に「公表」とは

⇒広く一般に自己の意思を知らせること（不特定多数の人々が知ることができるように発表すること）をいい、公表に当たっては、事業の性質及び個人情報の取扱状況に応じ、合理的かつ適切な方法によらなければならない。

➤ 本人の「同意」とは

⇒本人の個人情報が、個人情報取扱事業者によって示された取扱方法で取り扱われることを承諾する旨の当該本人の意思表示をいう（当該本人であることを確認できていることが前提となる）。

なお、個人情報の取扱いに関して同意したことによって生ずる結果について、未成年者、成年被後見人、被保佐人及び被補助人が判断できる能力を有していないなどの場合は、親権者や法定代理人等から同意を得ること。

12

3.事業者が守るべきルール ① - 取得・利用

【個人情報を取得する前】

1. 個人情報の**利用目的を特定**する。(第15条)

【個人情報を取得するとき】

1. **利用目的を予め公表**するか、取得する時に本人に**通知または公表**する。(第18条)
(書面での取得の場合は、利用目的を本人に**明示**する。)
2. **要配慮個人情報**を取得する場合は、原則本人の**同意**を得る。(第17条)
⇒詳細は次ページを参照
3. 不正の手段による取得はしない。(第17条)

【個人情報を利用するとき】

1. 利用目的の範囲内で利用する。(第16条)
2. **利用目的の範囲を超える場合は、本人の同意**を得る。(第16条)
3. 利用目的を変更することもできる。(関連性があると合理的に認められる場合に限り) (第15条)
その際は変更された利用目的を本人に**通知または公表**する。(第18条)
※法改正により当初の利用目的から新たな利用目的への変更の要件が緩和された

【個人情報の取扱いに関する苦情を受けたとき】

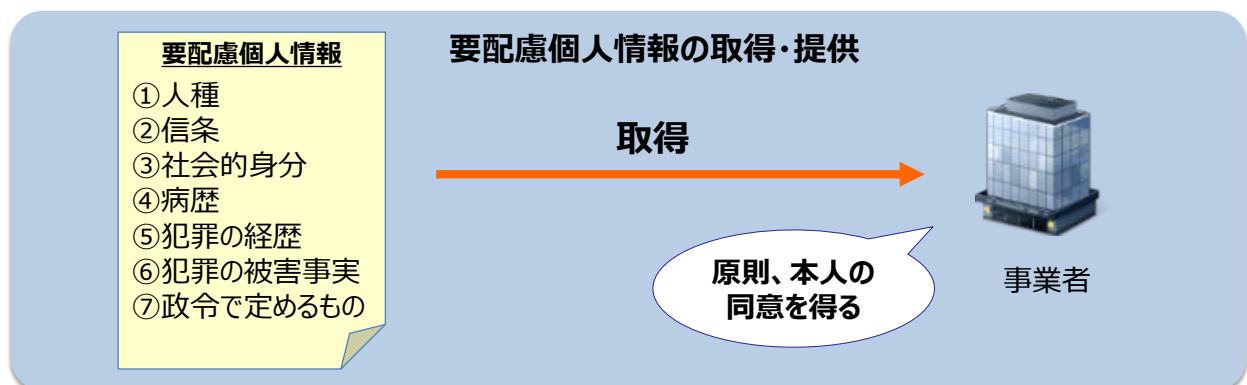
適切かつ迅速な処理に努める。(第35条)

13

3.事業者が守るべきルール ① - 取得・利用

○要配慮個人情報

本人に対する不当な差別、偏見が生じないように特に配慮を要するものとして、「要配慮個人情報」の定義を新設し、**取得、第三者提供**する際には、原則として本人の**同意**を得ることが必要となった。



⑦の政令の内容（政令第2条、規則第5条）

- ・身体障害・知的障害・精神障害等があること
- ・健康診断その他の検査の結果
- ・保健指導・診療・調剤情報
- ・本人を被疑者又は被告人として、逮捕、搜索等の刑事事件に関する手続が行われたこと
- ・本人を非行少年又はその疑いのある者として、保護処分等の少年の保護事件に関する手続が行われたこと

14

3.事業者が守るべきルール

個人データ

個人情報データベース等を構成する個人情報



保有個人データ

その事業者に開示等の権限のある個人データ (6ヶ月以内に消去するものを除く)



15

3.事業者が守るべきルール ② - 安全管理措置

【個人データを保管・管理するとき】

1. 正確で最新の内容に保ち、必要なくなったときはデータを消去するよう努める。(第19条)
 2. 安全管理措置を講じる。(第20条)
 3. 従業員に対して、必要かつ適切な監督を行う。(第21条)
 4. 委託する場合、委託先に対して必要かつ適切な監督を行う。(第22条)
- 法改正により、**取り扱う個人情報の数が5000人以下である事業者を規制の対象外とする制度を廃止**
 - ガイドラインでは、安全管理措置について、一般的な義務・手法例とは別に、**小規模の事業者**においても履行し得るような手法例を示している (通則編P86～P98)

※ガイドラインにおける「小規模の事業者」とは、

- 従業員の数が100人以下の事業者であって、次に掲げる事業者を除く事業者

① 取り扱う個人情報の数が5,000人分超の事業者

② 委託に基づいて個人データを取り扱う事業者

※安全管理措置として求められる要素の例

「取扱の基本的なルールを決める」、「従業員を教育する」、
「関係者以外が個人データを見れないようにする (漏えい防止含む。）」、
「PC等を用いて利用する場合はセキュリティ対策ソフトウェア等を導入する」等

16

3.事業者が守るべきルール ② - 安全管理措置

！ 小規模事業者向けの安全管理措置の手法例とヒント①

講じなければならない措置	手法例	ヒント💡
1 基本方針の策定	※この項目は、義務ではありません。	・義務ではありませんが、策定しておくことで、従業員教育に役立ちます。
2 個人データの取扱いに係る規律の整備	個人データの取得、利用、保存等を行う場合の基本的な取扱方法を整備する。	・既存の業務マニュアル・チェックリスト・フローチャート等に個人情報の取扱いの項目を入れるのも一案。
3 組織的安全管理措置		
(1) 組織体制の整備	個人データを取り扱う従業員が複数いる場合、責任ある立場の者とその他の者を区分する。	・リスク分散の観点から、誰かがチェックできると良いということです。
(2) 個人データの取扱いに係る規律に従った運用	あらかじめ整備された基本的な取扱方法に従って個人データが取り扱われていることを、責任ある立場の者が確認する。	・業務日誌やチェックリスト等を活用し、確認を。
(3) 個人データの取扱状況を確認する手段の整備		
(4) 漏えい等の事案に対応する体制の整備	漏えい等の事案の発生時に備え、従業員から責任ある立場の者に対する報告連絡体制等をあらかじめ確認する。	・「ほう・れん・そう」の中に、個人情報の漏えい事案を。
(5) 取扱状況の把握及び安全管理措置の見直し	責任ある立場の者が、個人データの取扱状況について、定期的に確認を行う。	・(1)～(4)のプロセスで気づいたりリスクがあれば、改善を。

17

3.事業者が守るべきルール ② - 安全管理措置

！ 小規模事業者向けの安全管理措置の手法例とヒント②

講じなければならない措置	手法例	ヒント💡
4 人的安全管理措置		
従業員の教育	- 個人データの取扱いに関する留意事項について、従業員に定期的な研修等を行う。 - 個人データについての秘密保持に関する事項を就業規則等に盛り込む。	・集合研修に限らず、朝礼等の際に定期的に注意喚起を。
5 物理的安全管理措置		
(1) 個人データを取り扱う区域の管理	個人データを取り扱うことのできる従業員及び本人以外が容易に個人データを閲覧等できないような措置を講ずる。	・誰でも見られる場所に放置しない。
(2) 機器及び電子媒体等の盗難等の防止	- 個人データを取り扱う機器、個人データが記録された電子媒体又は個人データが記載された書類等を、施錠できるキャビネット・書庫等に保管する - 個人データを取り扱う情報システムが機器のみで運用されている場合は、当該機器をセキュリティワイヤー等により固定する。	・書類や電子媒体をきちんと管理。
(3) 電子媒体等を持ち運ぶ場合の漏えい等の防止	個人データが記録された電子媒体又は個人データが記載された書類等を持ち運ぶ場合、パスワードの設定、封筒に封入し鞆に入れて搬送する等、紛失・盗難等を防ぐための安全な方策を講ずる。	・電子媒体にはパスワードを。置き忘れ等にも注意を。

18

3.事業者が守るべきルール ② - 安全管理措置

！ 小規模事業者向けの安全管理措置の手法例とヒント③

講じなければならない措置	手法例	ヒント💡
5 物理的安全管理措置		
(4) 個人データの削除及び機器、電子媒体等の廃棄	個人データを削除し、又は、個人データが記録された機器、電子媒体等を廃棄したことを、責任ある立場の者が確認する。	書類であれば、焼却、シュレッダー処理を、機器・電子媒体等であれば、データ削除ソフトウェアの利用や物理的な破壊等を。
6 技術的安全管理措置		
(1) アクセス制御	個人データを取り扱うことのできる機器及び当該機器を取り扱う従業者を明確化し、個人データへの不要なアクセスを防止する。	必要のない者の個人情報へのアクセスを制限するため、個人情報を含むファイルにパスワードを。
(2) アクセス者の識別と認証	機器に標準装備されているユーザー制御機能（ユーザーアカウント制御）により、個人情報データベース等を取り扱う情報システムを使用する従業者を識別・認証する。	
(3) 外部からの不正アクセス等の防止	- 個人データを取り扱う機器等のオペレーティングシステムを最新の状態に保持する。 - 個人データを取り扱う機器等にセキュリティ対策ソフトウェア等を導入し、自動更新機能等の活用により、これを最新状態とする。	セキュリティ対策ソフトウェアを最新の状態に。
(4) 情報システムの使用に伴う漏えい等の防止	メール等により個人データの含まれるファイルを送信する場合に、当該ファイルへのパスワードを設定する。	それほど難しい操作ではないので、メール送信時にはパスワードを。

19

3.事業者が守るべきルール ② - 安全管理措置

○事業者において個人データの漏えい等の事案が発生した場合等の対応（概要）

対象事案

- ✓ 個人データ（特定個人情報に係るものを除く。）の漏えい、滅失又は毀損
- ✓ 加工方法等情報（匿名加工情報の加工の方法に関する情報等）の漏えい
- ✓ これらのおそれ

望ましい対応

- (1) 事業者内部における報告及び被害の拡大防止
- (2) 事実関係の調査及び原因の究明
- (3) 影響範囲の特定
- (4) 再発防止策の検討及び実施
- (5) 影響を受ける可能性のある本人への連絡（事案に応じて）
- (6) 事実関係及び再発防止策等の公表（事案に応じて）



努力義務

個人情報保護委員会等への速やかな報告

※事業分野によっては、認定個人情報保護団体や権限の委任を受けた事業所管大臣が報告先となることがあります。
 ※なお、別途、業法等で監督当局への報告が義務付けられている場合もある為、注意が必要です。

20

3.事業者が守るべきルール ③ - 第三者に渡す場合

- **個人情報**を第三者に提供する時は、原則として**本人の同意が必要**。(第23条)

例外：①法令に基づく場合

②人の生命、身体又は財産の保護のため（かつ本人の同意を得ることが困難）

③公衆衛生・児童の健全な育成のため（かつ本人の同意を得ることが困難）

④国や地方公共団体等への協力

- 本人の同意を得ない場合には、以下(1)～(3)の手続をする（いわゆるオプトアウト手続）。ただし、**要配慮個人情報については、この手続による提供は禁止**。(第23条2項)

(1)本人の求めに応じて、その本人のデータの提供を停止することとする。

(2)以下の①～⑤をHPに掲載するなど、本人が容易に知ることができる状態にしておく。

①第三者提供を利用目的としていること、②提供される個人データの項目、③提供の方法、④本人の求めに応じて提供を停止すること、⑤本人の求めを受け付ける方法

(3)**本人に通知した事項を個人情報保護委員会に届け出る（個人情報保護委員会**
はこれを公表する）。

- 業務の委託、事業の承継、共同利用は、第三者提供には当たらない。(第23条5項)

- **第三者へ提供した時は、受領者の氏名等を記録し、一定期間保存する**。(第25条)

- **第三者から個人データを受け取る時は、提供者の氏名等、取得経緯を確認し、受領年月日、確認した事項等を記録し、一定期間保存する**。(第26条)

⇒詳細は次ページ参照

21

3.事業者が守るべきルール ③ - 第三者に渡す場合

○記録義務に関する委員会規則（規則第12条～18条）の内容

本規定は名簿屋対策が目的のため、一般的なビジネスの実態に配慮して次のとおり整理する。

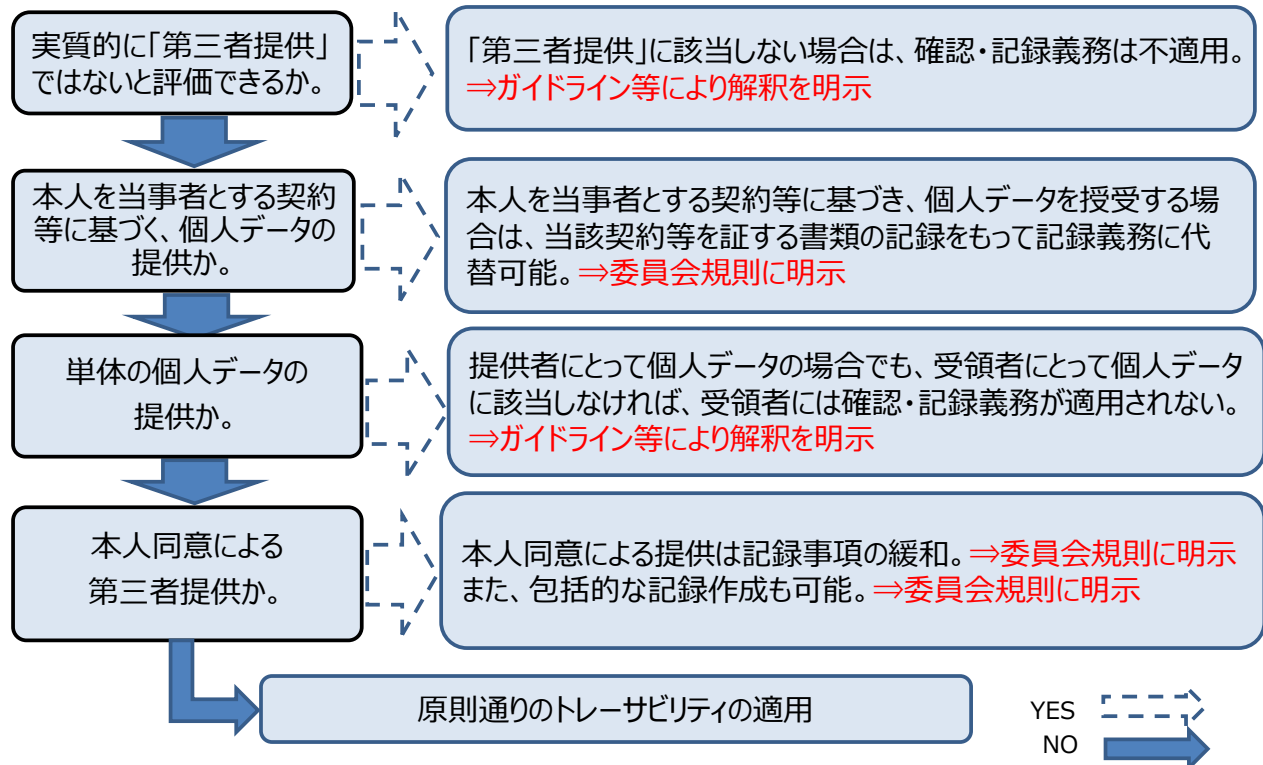
- 記録事項として、第三者提供について本人同意がある場合は、提供年月日の記録は不要とする。
- 記録の保存期間については、原則3年とするが、本人に対する物品等の提供に関連して本人同意のもとで第三者提供した場合は1年とする。
- 本人との契約等に基づく提供については、既存の契約書等で代替可能とする。
- 反復継続して提供する場合は包括的な記録で足りることとする。

ガイドラインでは、一般的なビジネスの実態に配慮して、次のようなケースでは確認・記録義務がかからないと整理（解釈で対応）

- ・本人による提供と整理できるケース（例：SNS上の個人のプロフィール）
- ・本人に代わって提供と整理できるケース（例：銀行振込）
- ・本人側への提供と整理できるケース（例：同席している家族）
- ・「個人データ」に該当しないと整理できるケース（例：名刺1枚）等

22

3.事業者が守るべきルール ③ - 第三者に渡す場合



※平成28年7月29日付委員会資料「改正個人情報保護法第25条・第26条の確認・記録義務の方向性と委員会規則（案）の対応表」より抜粋

23

3.事業者が守るべきルール ④ - 外国の第三者に渡す場合

○次の①～③のいずれかに該当する必要がある。（第24条）

- ① 外国にある第三者へ提供することについて、本人の同意を得る。
- ② 外国にある第三者が個人情報保護委員会の規則で定める基準に適合する体制を整備している。
- ③ 外国にある第三者が個人情報保護委員会が認めた国に所在する。

外国への第三者提供に関する規則（規則第11条）の内容

②の「個人情報保護委員会の規則で定める基準適合する体制」について、一般的なビジネスの実態に配慮して次に該当するものと整理する。

- 提供を受ける者における個人データの取扱いについて、適切かつ合理的な方法により、個人情報保護法の趣旨に沿った措置の実施が確保されていること
- 個人データの提供を受ける者が、個人情報の取扱いに係る国際的な枠組みに基づく認定を受けていること

ガイドラインでは、規則で定められた基準について、具体的な事例も交えて分かりやすく示している。

- ・「適切かつ合理的な方法」の例：委託契約やグループ企業の内規・プライバシーポリシー等
- ・「個人情報保護法の趣旨に沿った措置」の具体例：OECD、APEC等の国際的な枠組みの基準に基づいて記載
- ・「個人情報の取扱いに係る国際的な枠組み」：「APECの越境プライバシールール（CBPR）システム」を記載

24

3.事業者が守るべきルール ⑤ - 開示請求への対応

【保有個人データに対する開示請求等への対応】

1. ①事業者の名称
②利用目的
③請求手続
④苦情申出先、
⑤加入している認定団体個人情報保護団体の名称・苦情申出先
を本人の知り得る状態に置く。（第27条）
2. 本人から開示を**請求**された場合、原則本人に開示する。（第28条）
3. 本人からの**請求**に応じて、
 - ①内容に誤りがある場合には、訂正・削除をする。（第29条）
 - ②利用目的制限（第16条）、適正取得（第17条）違反の場合、利用停止・消去する。第三者提供の制限（第23条1項）、外国の第三者への提供制限（第24条）違反の場合、第三者への提供を停止する。（第30条）

※改正前は保有個人データの開示等の「求め」は、「裁判所に訴えを提起することができる権利（請求権）」であると明確にされていなかったが、**法改正により「請求権」であることが明確化された**

25

3.事業者が守るべきルール - 罰則

●国の監督（第40条～46条）

国は事業者に対して、必要に応じて報告を求めたり**立入検査**を行うことができる。
また、実態に応じて、**指導・助言、勧告・命令**を行うことができる。

※法改正により**個人情報保護委員会に監督権限が一元化され、立入検査、指導の権限が追加された**

●罰則

○国からの命令に違反した場合（第84条）
⇒6ヶ月以下の懲役又は30万円以下の罰金

○虚偽の報告等をした場合（第85条）
⇒30万円以下の罰金

○従業員等が不正な利益を図る目的で**個人情報データベース等を提供、又は、盗用した場合（個人情報データベース等不正提供罪）**（第83条）
⇒1年以下の懲役又は50万円以下の罰金

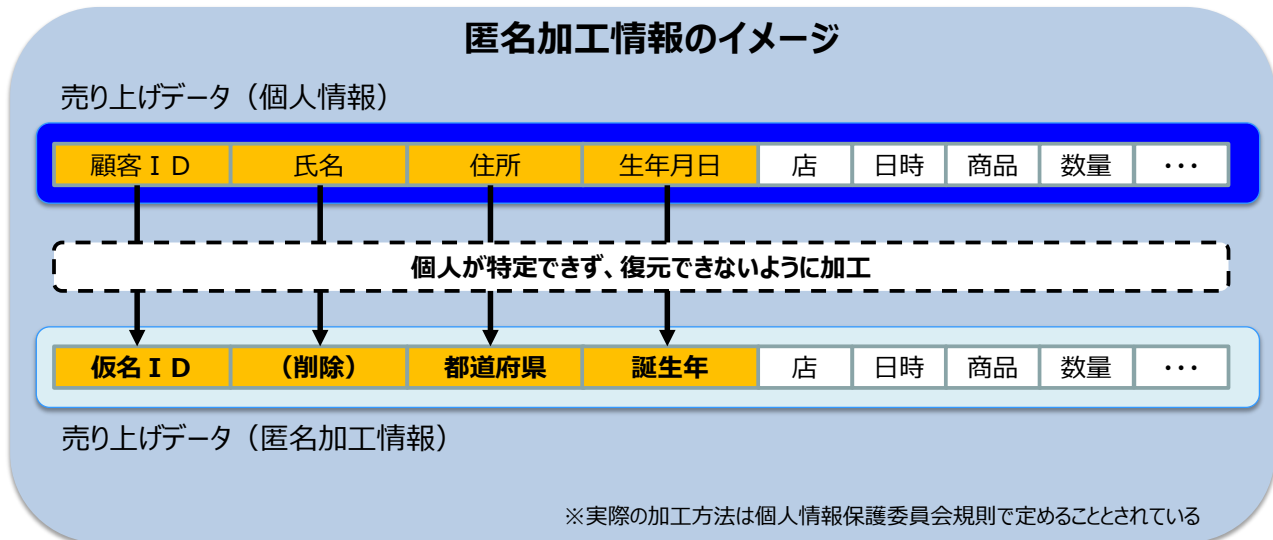
※改正前は、不正な利益を図る目的で個人情報を提供した者に対して、その個人に適用できる罰則規定がなかったが、改正法では、**新しく「個人情報データベース等不正提供罪」が設けられた**

26

4.匿名加工情報とは

○匿名加工情報の概要

個人情報を**特定の個人を識別することができず**、かつ、当該個人情報を**復元することができないようにしたものを「匿名加工情報」と定義し**、本人の同意に代わる一定の条件の下、自由に利活用できる環境が整備された。



27

4.匿名加工情報とは

○匿名加工情報取扱事業者等の義務

【匿名加工情報を作成するとき】（第36条）

- ①適正な加工を行う。
- ②安全管理措置を講じる。
- ③匿名加工情報に含まれる情報の項目を公表する。
- ④加工前の個人情報における本人の特定をしようとししない。
- ⑤苦情の処理等に努める。

【匿名加工情報を第三者へ提供するとき】（第36条、第37条）

- ①提供する匿名加工情報に含まれる「情報の項目」と「提供の方法」を公表する。
- ②提供先に対して匿名加工情報であることを明示する。

【匿名加工情報を第三者から受領するとき】（第38条、第39条）

- ①加工前の個人情報における本人の特定をしようとししない。
- ②苦情の処理等に努める。

28

4.匿名加工情報とは

「匿名加工情報」に関する規則の内容

匿名加工情報の作成方法に関して、最低限の規律として、次の措置を講ずることを求める。
なお、詳細は自主ルールに委ねる。

- ① **特定の個人を識別することができる記述等（例：氏名）の全部又は一部を削除**（置換を含む。以下同じ。）すること
- ② **個人識別符号の全部を削除**すること
- ③ 個人情報と他の情報とを**連結する符号**（例：委託先に渡すために分割したデータとひも付けるID）を削除すること
- ④ **特異な記述等（例：年齢116歳）を削除**すること
- ⑤ 上記のほか、個人情報とデータベース内の他の個人情報との差異等の**性質を勘案し、適切な措置**を講ずること

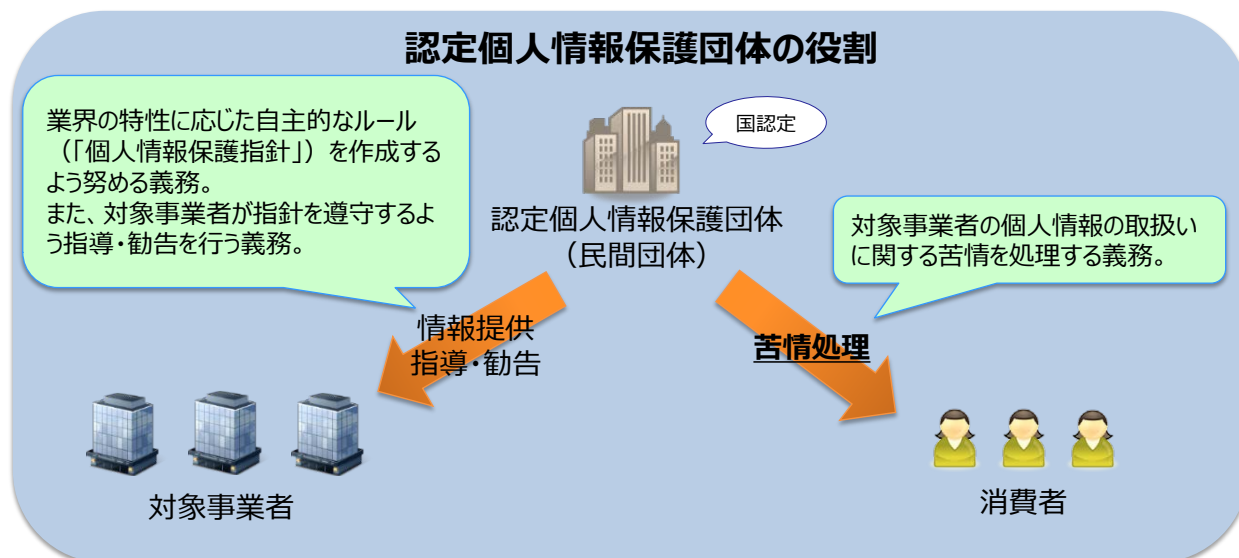
ガイドラインでは、規則第19条で定められた匿名加工情報の作成方法に関する上記の基準等について、具体的な事例等も交えて、分かりやすく示している。
その他、匿名加工の手法、データ処理等について、認定個人情報保護団体の自主ルールを作成する際の参考となる事項、考え方について示す事務局レポートも作成している。

29

5.お役立ち情報 ① - 認定個人情報保護団体

○認定個人情報保護団体の制度（第47条～58条）

- 事業者の個人情報の適切な取扱いの確保を目的として、国の認定を受けた民間団体。
- 対象事業者への情報提供、個人情報に関する苦情の処理等を行う。



30

5.お役立ち情報 ② - 適用除外

○個人情報保護法の適用が除外される事業者（第76条）

個人情報取扱事業者のうち、次に掲げる者が、それぞれ定められた目的で、個人情報等を取り扱う場合は、法の適用除外とされている。

- ① **放送機関、新聞社、通信社その他の報道機関**
⇒ **報道**の用に供する目的
- ② **著述を業として行う者**
⇒ **著述**の用に供する目的
- ③ **大学その他の学術研究を目的とする機関若しくは団体又はそれらに属する者**
⇒ **学術研究**の用に供する目的
- ④ **宗教団体**
⇒ **宗教活動**の用に供する目的
- ⑤ **政治団体**
⇒ **政治活動**の用に供する目的

31

5.お役立ち情報 ③ - お問合せ先

○苦情・相談のお問合せ先

● 個人情報保護法に関する質問等

個人情報保護法の解釈についての一般的な質問や、事業者における取扱いに関する苦情等は下記にご連絡ください。

個人情報保護法相談ダイヤル

03-6457-9849
くわしく

受付時間 土日祝日及び年末年始を除く 9：30～17：30

●事業者の個人情報の取扱いに関する苦情相談

事業者の個人情報の取扱いに関する苦情相談は、以下の窓口にお問合せください。

- 事業者の苦情受付窓口
- 消費生活センター等の地方公共団体の窓口
- 認定個人情報保護団体 など

32

5.お役立ち情報 ③ - お問い合わせ先

マイナンバー苦情あっせん相談窓口

個人情報保護委員会では、マイナンバー(個人番号)の取扱いに関する苦情の申出についての必要あっせんを行うため、電話による苦情あっせん相談窓口を設置しています。
マイナンバー(個人番号)の取扱いに関する苦情をお持ちの方なら、どなたでもご利用いただけます。

03-6457-9585

土日祝日及び年末年始を除く 9:30～17:30

○ 相談窓口にてできること

- 1 苦情の内容を所掌する他の相談窓口の紹介
- 2 苦情の相手方への苦情の内容の伝達(委員会が必要と認めた場合)
- 3 番号法に定められた措置等に反する行為があった場合の監督部門への取次ぎ
- 4 苦情をめぐって、苦情の相手方と争いが生じた場合のあっせん

○ 相談例

- 1 事業者に苦情を申し立てたが、対応してもらえない。
- 2 事業者の苦情に対する対応に不満があるが、どうしたらよいか分からない。
- 3 事業者でマイナンバー(個人番号)が漏えいしており、自分の情報が流出している可能性がある。
- 4 番号法で定められた措置がなされず、自分の情報が適正に管理されていない。
- 5 ある事業者でマイナンバー(個人番号)に関する不適切な処理がなされている。



愛称:マイナちゃん

33

5.お役立ち情報 ④ - 参考資料

○個人情報保護法に関する参考資料

本資料は、個人情報保護法の概要をまとめたものであり、事業者の義務や例外規定の全てを記載したものではありません。

個人情報保護法の解釈やより詳細な内容については、以下のウェブページをご参照下さい。

- 個人情報保護委員会ホームページ
<http://www.ppc.go.jp/>
- ガイドライン等
<https://www.ppc.go.jp/personal/legal/>
- 中小企業サポートページ(個人情報保護法)
http://www.ppc.go.jp/personal/chusho_support/
- 会員名簿を作るときの注意事項(自治会・同窓会向け)
http://www.ppc.go.jp/files/pdf/meibo_sakusei.pdf

34

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について

内閣官房 健康・医療戦略室

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の概要 (次世代医療基盤法)

法律の目的

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し、匿名加工医療情報作成事業を行う者の認定、医療情報及び匿名加工医療情報等の取扱いに関する規制等を定めることにより、健康・医療に関する先端的研究開発及び新産業創出を促進し、もって健康長寿社会の形成に資することを目的とする。

法律の内容

1. 基本方針の策定

政府は、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施策の推進を図るための基本方針を定める。

2. 認定匿名加工医療情報作成事業者(以下「認定事業者」という。)

主務大臣は、申請に基づき、匿名加工医療情報作成事業の適正かつ確実な実施に関する基準に適合する者を認定する。

①認定事業者の責務

- ・医療情報の取扱いを認定事業の目的の達成に必要な範囲に制限する。
- ・医療情報等の漏えい等の防止のための安全管理措置を講じる。
- ・従業者に守秘義務(罰則付き)を課す。
- ・医療情報等の取扱いの委託は、主務大臣の認定を受けた者に対してのみ可能とする。

②認定事業者の監督

- ・主務大臣は、認定事業者に対して必要な報告徴収、是正命令、認定の取消し等を行うことができる。

3. 認定事業者に対する医療情報の提供

医療機関等は、あらかじめ本人に通知し、本人が提供を拒否しない場合、認定事業者に対し、医療情報を提供することができる。(医療機関等から認定事業者への医療情報の提供は任意)

4. その他

主務大臣は、内閣総理大臣、文部科学大臣、厚生労働大臣及び経済産業大臣とする(認定事業者の認定等については、個人情報保護委員会に協議する)。

※生存する個人に関する情報に加え、死亡した個人に関する情報も保護の対象とする。

施行期日

公布の日から起算して1年を超えない範囲内において政令で定める日

次世代医療基盤法によって実現できること(例)

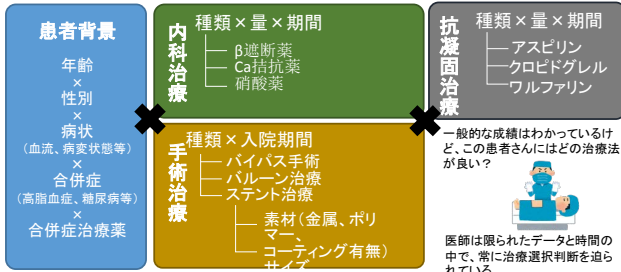
自らが受けた治療や保健指導の内容や結果を、データとして研究・分析のために提供し、その成果が自らを含む患者・国民全体のメリットとして還元されることへの患者・国民の期待にも応え、ICTの技術革新を利用した治療の効果や効率性等に関する大規模な研究を通じて、患者に最適な医療の提供を実現する。

■ 治療効果や評価等に関する大規模な研究の実現

例1) 最適医療の提供

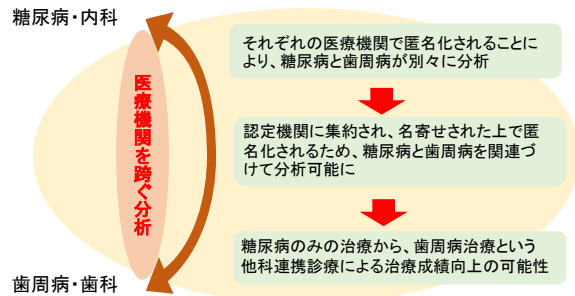
・大量の実診療データにより治療選択枝の評価等に関する大規模な研究の実施が可能になる。

<例: 狭心症治療>



例2) 異なる医療機関や領域の情報を統合した治療成績の評価

・糖尿病と歯周病のように、別々の診療科の関連が明らかになり、糖尿病患者に対する歯周病治療が行われることで、健康状態が向上する可能性



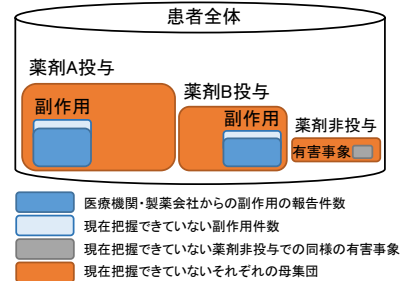
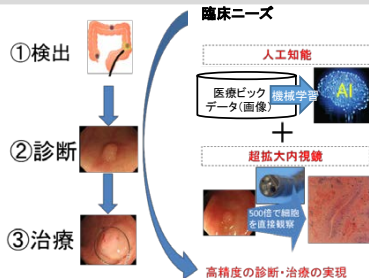
■ 医薬品市販後調査等の高度化、効率化

<医薬品等の安全対策の向上>

・副作用の発生頻度の把握や比較が可能になり、医薬品等の使用における更なる安全性の向上が可能に

例3) 最先端の診療支援ソフトの開発

・人工知能(AI)も活用して画像データを分析し、医師の診断から治療までを包括的に支援

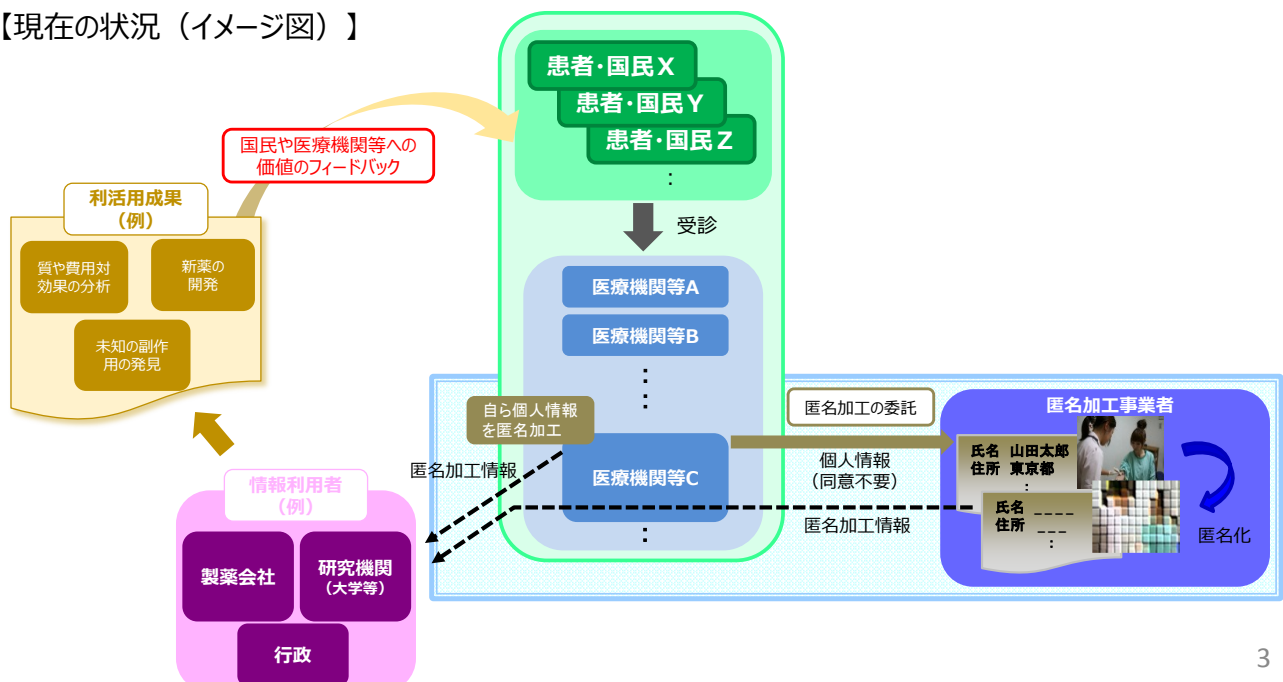


2

現行法で可能な匿名加工医療情報の提供の仕組み

- 匿名加工情報とは、特定の個人を識別することができないよう個人情報加工して得られる個人に関する情報であって、当該個人情報を復元することができないようにしたもの。
- 匿名加工情報については、本人の同意なく第三者に対する提供が可能。
- このため、個別医療機関は、保有する医療情報（個人情報）の匿名加工を自ら又は事業者へ委託して行い、利活用者に本人の同意なく提供することは可能である。

【現在の状況（イメージ図）】



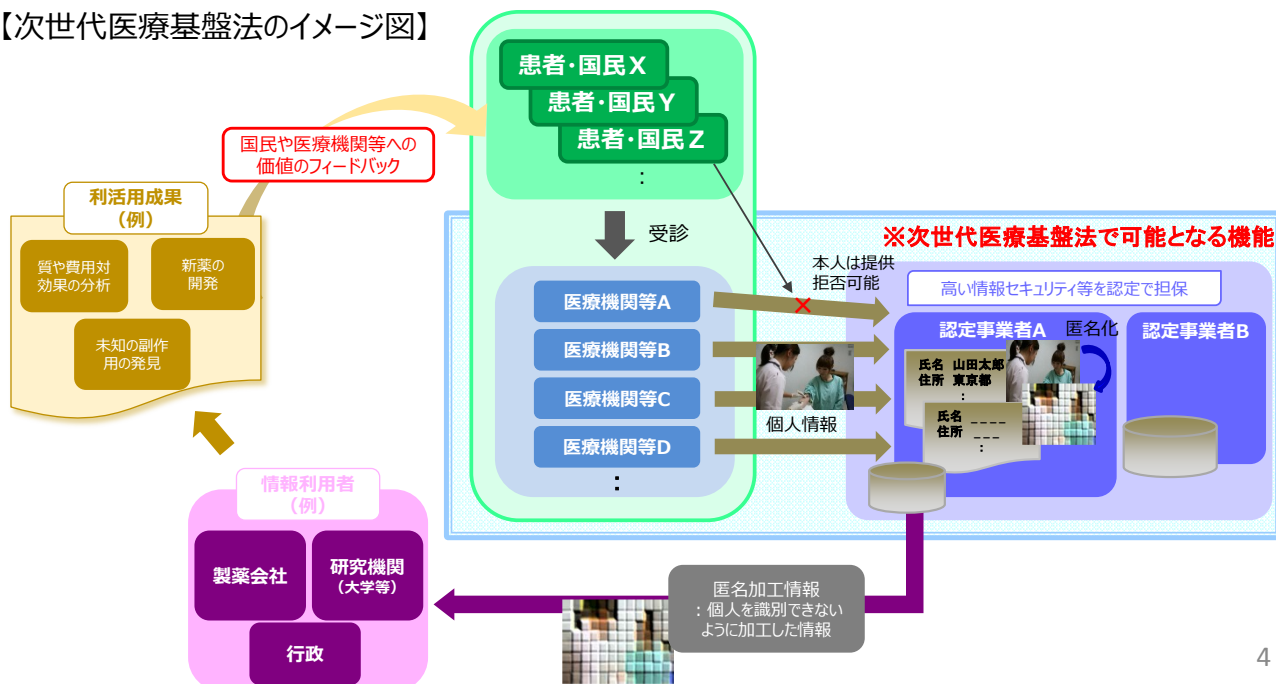
3

次世代医療基盤法の全体像(匿名加工医療情報の円滑かつ公正な利活用の仕組みの整備)

個人の権利利益の保護に配慮しつつ、匿名加工された医療情報を安心して円滑に利活用することが可能な仕組みを整備。

- ①高い情報セキュリティを確保し、十分な匿名加工技術を有するなどの一定の基準を満たし、医療情報の管理や利活用のための匿名化を適正かつ確実に行うことができる者を認定する仕組み(=認定匿名加工医療情報作成事業者)を設ける。
- ②医療機関等は、本人が提供を拒否しない場合、認定事業者に対し、医療情報を提供できることとする。
認定事業者は、収集情報を匿名加工し、医療分野の研究開発の用に供する。

【次世代医療基盤法のイメージ図】



4

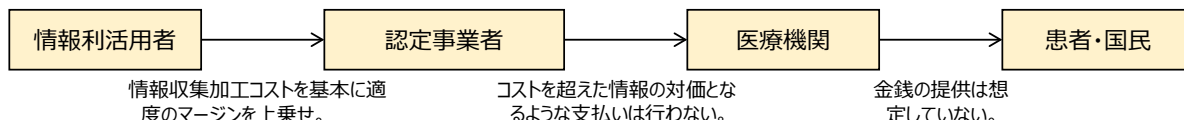
次世代医療基盤法のポイント

1 目的、目指すもの

- 次世代医療基盤法は、医療情報について特定の個人を識別できないよう匿名加工する事業者に対する規制を整備し、匿名加工情報の安心・適正な利活用を通じて、健康寿命の延伸、健康長寿社会の実現を目指すもの。
- これにより、治療効果等に関する大規模な研究を通じた最適な医療の提供や医薬品副作用等の早期把握による安全性の向上等を患者・国民へ還元する効果が期待できる。

2 利活用の基本的な考え方

- 匿名加工された医療情報の円滑かつ公正な利活用を、基本方針（閣議決定）と認定基準で確保。
- (1)利活用の基本方針を閣議決定で明確化。認定事業者は、情報利活用者の目的を個別に確認し、健康長寿社会の実現に資する公益性の高い研究開発に優先的に対応。不適切な利用は排除し、意図せざる目的への転用を防止。
- (2)情報の収集加工提供に要するコストの利活用者への転嫁を基本とし、認定事業者に過度な利潤を生じさせない。



3 セキュリティ

- 事前及び万一の場合の事後の対応を徹底。
具体的には、①組織・人的要因によるリスク排除、②基幹システムのオープンネットワークからの分離、③多層防壁・安全策の導入（個人情報の暗号化や緊急時の対応・監督体制の確保を含む。）を徹底。

4 段階的運用

- 認定事業者による匿名加工事業は、適正な規模で始め、成果等を確認しつつ徐々に拡大。

5

認定匿名加工医療情報作成事業者のセキュリティ確保の基本的考え方

安全面での課題	情報の漏洩 盗み見 情報・システムの改変・破壊	被害	個人の医療情報の悪用 誤情報の活用、業務停止 認定匿名加工医療情報作成事業者への信頼喪失等
---------	-------------------------------	----	---

基本的手口(複数の組合せによる)

- ① 騙し・なりすましによる暗証等の入手
- ② 標的型攻撃メール等によるネットワークからの侵入・操作
- ③ ソフトウェアの脆弱性の利用、不正通信ソフトウェア、ハードウェアの製造工程における意図せざる変更
- ④ 内部の不正アクセス(盗み見、記録メディアによる情報の持出し)

対応方針

- ① 組織・人的要因の徹底排除
- ② 基幹システムはオープンネットワークから分離
- ③ 多層防御・安全策の導入(想定外の手口にも対応)

具体策(「三本の柱」)

① 組織・人的要因の徹底排除

- ・教育・運用・管理体制の整備(罰則付守秘義務)
- ・警備員・監視カメラ・入退室管理

② 基幹システムはオープンネットワークから分離

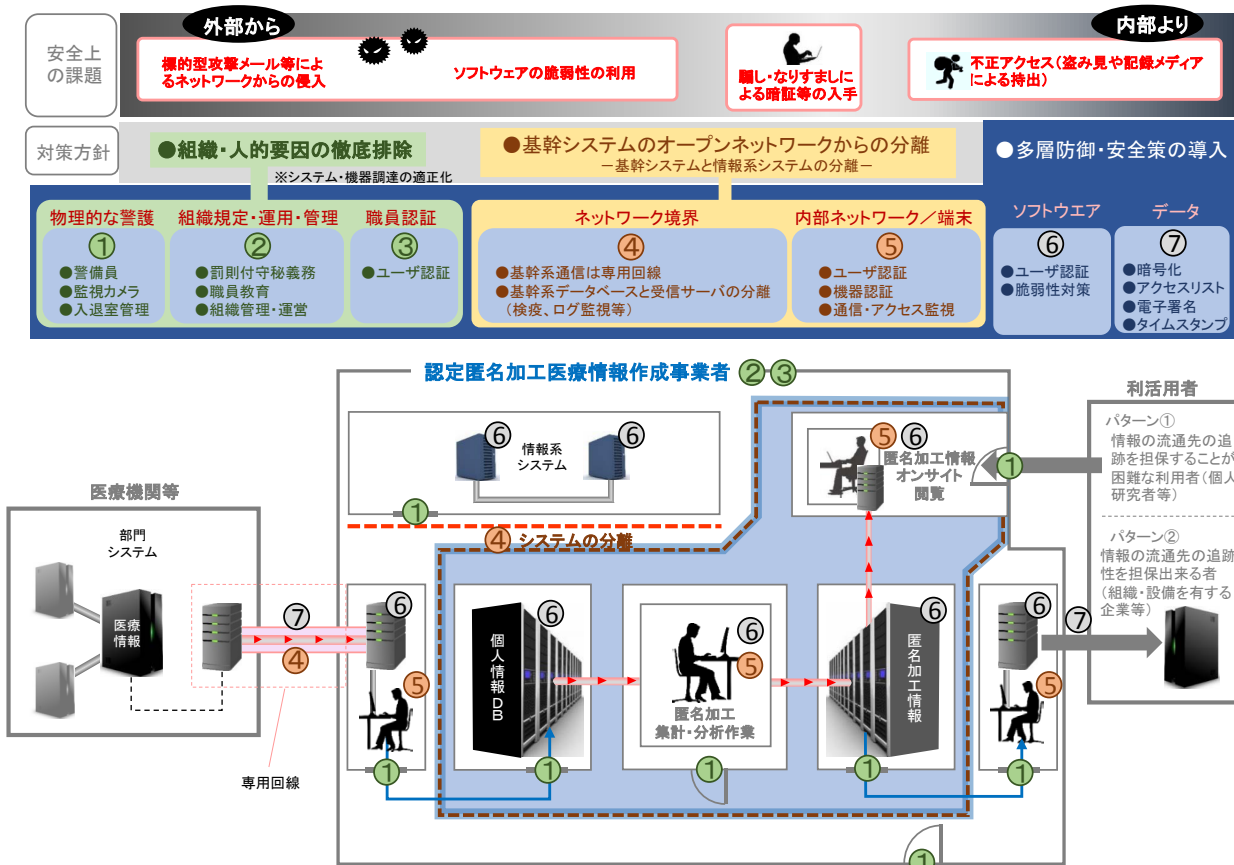
- ・基幹業務系と情報系システムの分離
- ・基幹業務系はインターネット等オープン環境から分離

③ 多層防御・安全策の導入(想定外の手口にも対応)

- ・基幹業務に係るデータの送受信は、基幹業務データベースと切り離し実施(ファイアーウォール等)
(それぞれ対応状況の異なる医療機関のセキュリティ水準に影響を受けないよう認定事業者の責任においてセキュリティ対策を実施)
- ・アクセスログ/データ操作ログをリアルタイムで監視(予定されない通信、アクセスは直ちに遮断する等)
- ・記録メディアの制限
・ソフトウェアの不断のアップデート(脆弱性対応等)
- ・データの暗号化(万が一、悪意ある者がデータ断片を入手しても解読困難)
- ・匿名加工情報利用者側のデータ利用の追跡性(トレーサビリティ)確保
- ・第三者認証を含む継続的なセキュリティ水準の確保や緊急時の対応、監督官庁への連絡体制の確保

6

認定匿名加工医療情報作成事業者の具体的セキュリティー対策(詳細版)



7

匿名加工について

- 匿名加工情報とは、特定の個人を識別することができないよう個人情報加工して得られる個人に関する情報であって、当該個人情報復元することができないようにしたもの（「一般人をもって具体的な人物と情報の間に同一性を認めるに至ることができるか」により判断される）
- 個人情報保護委員会が定める基準（①特定個人を識別可能な記述、②個人識別符号、③個人情報の連結符号、④特異な記述を削除する、または①～④を復元できない方法により他の記述等に置き換える）に沿って適確に匿名加工を行う能力を有する事業者を認定する。

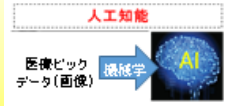
（※）匿名加工情報については、本人の同意なく第三者に対する提供が可能。（匿名加工情報については、本人を識別するための照合等を禁止）

- 匿名加工情報の提供範囲が無限定に拡散しないよう、認定匿名加工医療情報作成事業者では、利活用者との契約において、情報の共有範囲を明確化。

認定匿名加工医療情報作成事業者が行う医療分野の研究開発に資する匿名加工のイメージ（例）

<人工知能による診療支援システム>

- ① 人工知能による診療支援のために、大量の画像を機械学習させたい。
- 氏名、生年月日、性別等特定の個人を識別することができる記述を削除した上で、一般人をもって特定の個人の識別が不可能であるような画像は、匿名加工情報として提供することが可能。



<革新的な疫学研究>

- ② 複数の医療機関が保有する情報を個人別に突合し、市区町村別の集団毎の健康状態について分析したい。
- 認定事業者においてあらかじめ個人別に突合した上で、医療機関内での管理のために用いられているID等や、市区町村以下の住所情報や病院名を削除した匿名加工情報を提供可能



<医薬品市販後調査等の高度化、効率化>

- ③ 医薬品等の安全対策の向上のため、投薬等の医療行為と副作用等の発症の因果関係等の解析したい。
- 生年月日、投薬日等の日付情報を一律にずらすことにより、医療行為と副作用等の発生の関係を崩さずに情報を提供可能。



<臨床研究の高度化>

- ④ 治験の実施に当たり、軽症の糖尿病で、合併症がないような対象者等の分布をあらかじめ把握したい。
- 認定事業者内において必要な統計処理等をした結果を匿名加工情報又は統計情報として提供可能。



委員構成(敬称略)

座 長 橋爪 誠(九州大学大学院)

委 員 (五十音順)

佐々木 毅 (東京大学:日本病理学会推薦)
 清水 昭伸 (東京農工大学大学院)
 田中 聖人 (京都第二赤十字病院:日本消化器内視鏡学会推薦)
 中田 典生 (東京慈恵会医科大学:日本医用画像工学会推薦)
 藤田 広志 (岐阜大学:医用画像情報学会推薦)
 待鳥 詔洋 (国立国際医療研究センター 国府台病院:
 日本医学放射線学会推薦)
 横井 英人 (香川大学)

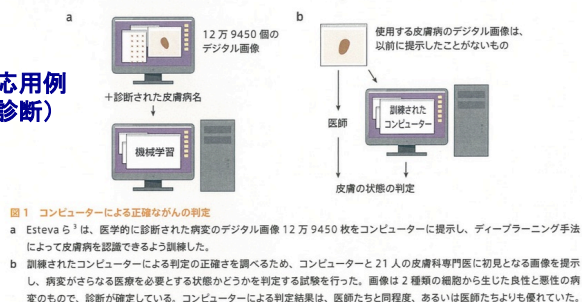
事務局(国立衛研)

葩島由二, 中岡竜介, 加藤玲子

(独)医薬品医療機器総合機構

医療機器審査第一部、第二部、第三部、規格基準部

座長含め
 医学系委員:6名
 工学系委員:2名
 うち学会推薦委員:5名
 (関連学会と連携)

人工知能応用例
(皮膚病診断)

Nature ダイジェスト, 14 (5), 29-31 (2017).

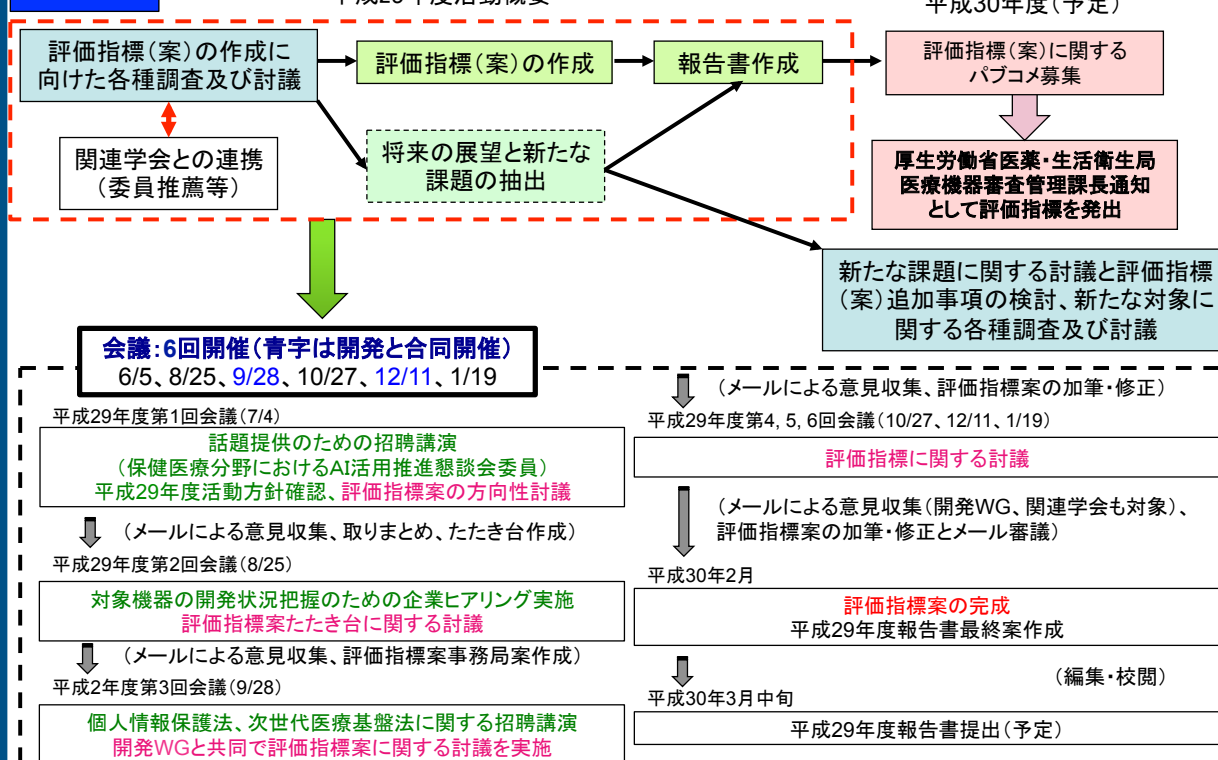
本審査WGの目標

コンピュータ診断支援を行う医療機器に関しては、既に本事業による次世代医療機器評価指標通知が平成23年12月7日に発出されているが、近年の計算機技術の著しい発展により、学習データのさらなる追加(事後学習)等により市販後も性能が変化する人工知能技術の適用はその対象とされていなかった。特に**深層学習**を用いた人工知能技術では、**最終結果を導き出すアルゴリズムがブラックボックス化**している点が特徴であると同時に、事後学習等による性能変化の内容も基本的にはアウトプットでしか評価できないこと等から、その評価や運用において従来にない**新たな課題**が生じることが懸念されている。当該診断支援装置を用いた場合における適切な性能、品質評価方法、特に市販後の性能変化に対する評価方法等に関して、**従来とは異なる新しい視点**も考慮する必要が生じた。このような背景に鑑み、本WGを立ち上げ、まず、**人工知能技術を利用した医用画像診断支援システム**の有効性と安全性を評価する際の問題点や留意すべき点について、現時点での考え方をとりまとめ、評価指標案を作成する。

平成29年度活動内容

活動概略

平成29年度活動概要

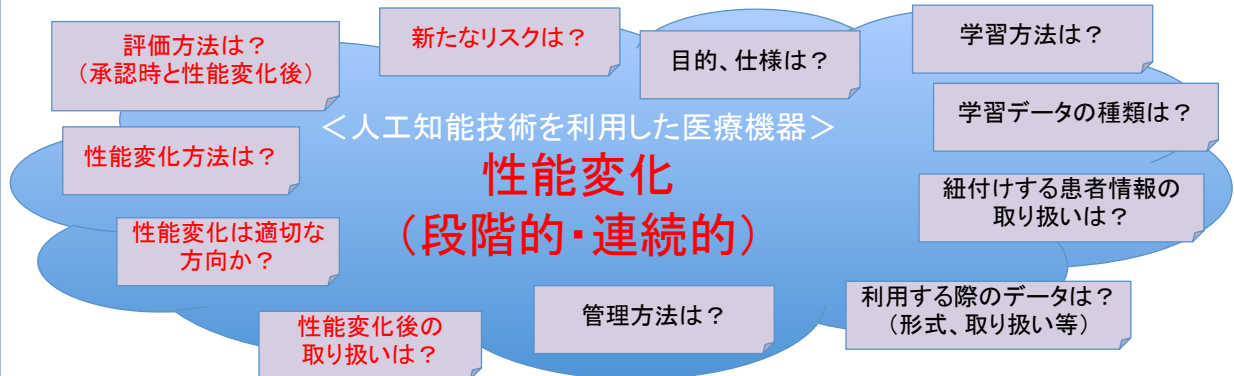


評価指標案の作成

目的

人工知能技術を利用した医用画像診断支援システムに対する包括的な評価指標(案)の作成

主な討議ポイント: その使用に伴って性能が変化する医療機器の取り扱い



- 性能が変化する支援システムの原理を示すことが難しい
(特に、深層学習の場合は診断アルゴリズムがブラックボックス化)
 - 必要に応じてプログラム概要を示すことを求めるに留め、適切に性能評価で品質を確保することが重要
- 変化した性能が承認申請時に規定された範囲を確保することが重要
 - 性能変化の範囲をどう考えるか(将来も考慮し、段階的変化と連続的変化の両方を想定)
 - 適切な性能検証方法をどのように考えるか(事後学習データの質は重要だが、想定外の変化は生じる)

評価指標案の作成とその最終構成、次年度計画

主な討議ポイント: 許容される性能変化の内容・範囲

- 承認申請時に性能範囲を規定する必要がある
 - ➔ 目的に応じて、臨床上許容される、あるいは統計学的な妥当性に基づいた範囲の設定が必要
 - ➔ 同時に適切な検証方法や、性能が逸脱した際の適切な対策も必要
- 個別システムに人工知能が搭載されている場合の取り扱い(システム毎に性能変化する場合)
 - ➔ 品質確保が難しい等、現時点では課題が多いため、本文ではなく別添として考え方を取りまとめた

構成

これらの討議内容を適宜反映させつつ評価指標案を取りまとめた

＜人工知能技術を利用した医用画像診断支援システムに関する評価指標案＞

1. はじめに
2. 用語の定義・説明
3. 本評価指標の対象
4. 本評価指標の位置付け
5. 支援システム等で想定される主な課題と解決の方向性
(1)ブラックボックスとしての性質 (2)性能変化 (3)責任の所在
6. 評価に際して留意すべき事項
7. 機器の特性に応じた追加留意事項

別添: エンドユーザ等が性能変化に関与可能な人工知能技術を利用した医用画像診断支援システムに関する基本的考え方

性能変化方式にかかわらず人工知能を
メーカーが管理している場合
例1) メーカーがバッチ方式で性能変化を実施
例2) リアルタイムで性能変化する人工知能
をメーカーが保有・管理

人工知能が個別の医療機器
に実装されており、臨床現場
毎でその性能変化が実施可
能な場合

次年度

- ・ 著しい当該技術発展に合わせた評価指標追加改訂の検討と別添内容の検討継続
- ・ 対象となる医療機器を拡大し、国内外における開発及び規制の実態調査を実施(新テーマの検討) 等

VI. 平成 29 年度 WG 会議議事概要

次世代医療機器・再生医療等製品評価指標作成事業
人工知能分野 審査 WG
平成 29 年度 第 1 回会議 議事概要

1. 開催日時：平成 29 年 6 月 5 日（月）16:00 – 19:10
2. 開催場所：TKP 品川カンファレンスセンター ANNEX カンファレンスルーム 7
（東京都港区高輪 3-13-1 TAKANAWA COURT 3F）
3. 出席者（委員：座長以下五十音順・敬称略）
座長：橋爪 誠（九州大学）
委員：佐々木 毅（東京大学）、清水 昭伸（東京農工大学）、田中 聖人（京都第二赤十字病院）、中田 典生（東京慈恵会医科大学）、藤田 広志（岐阜大学）、待鳥 詔洋（国立国際医療研究センター国府台病院）、横井 英人（香川大学）
厚生労働省（厚労省）：磯部 総一郎、青柳 ゆみ子
医薬品医療機器総合機構（PMDA）：牧野 勤、菅原 明彦、加藤 健太郎、佐々木 拓哉、桐山 瑤子、古森 亜矢、宮崎 生子、今川 邦樹
オブザーバー：松尾 豊（東京大学）、鎮西 清行（産業技術総合研究所（産総研））、坂無 英徳（産総研）、扇谷 悟（日本医療研究開発機構（AMED））、浅沼 直樹（AMED）
事務局（国立医薬品食品衛生研究所）：薮島 由二、中岡 竜介、植松 美幸
4. 配布資料
資料 1：第 1 回審査 WG 会議議事次第（案）
資料 2：委員名簿
資料 3：次世代医療機器・再生医療等製品評価指標作成事業概要
資料 4：PMDA 講演資料
資料 5：厚生労働省講演資料
資料 6：松尾先生講演資料
資料 7：評価指標案作成のために討議すべき点について（事務局案）
資料 8：メーリングリストについて
参考資料 1：「医療機器プログラムの承認申請に関するガイダンスの公表について」（平成 28 年 3 月 31 日付け事務連絡）
参考資料 2：「次世代医療機器評価指標の公表について」（平成 23 年 12 月 7 日付け薬食機発 1207 第 1 号通知）別添 3「コンピュータ診断支援装置に関する評価指標」
5. 議事内容
5-1. 開催挨拶及び資料確認
座長（橋爪委員）及び事務局（薮島）挨拶後、事務局より配布資料の確認・説明及び WG 委員を含めた全参加者が紹介された。

(1)活動計画
事務局より資料 3 を用いて次世代事業の趣旨及び会議の流れについて説明された。主な内容は以下のとおりである。

- ・ 目的は承認審査の迅速化、円滑化を目的とした評価指標案の作成である。
- ・ 今年度の対象は実用化が最も早いと見込まれている画像診断装置であり、平成 23 年 12 月 7 日付けで発出した「コンピュータ画像診断支援装置に関する評価指標」に AI 特有の要求事項を上乗せした評価指標案と考えていただいてもよい。
- ・ 「コンピュータ画像診断支援装置に関する評価指標」の発出後、クラウドコンピューティングや Deep Learning (DL) が急速に発展してきており、その利用において適切な考え方の補充を含め、包括的な検討が必要な時期にきている。また、AI を用いた医療機器の実用化も間近と考えられる。
- ・ AI の最大の特徴として性能の継続的変化が挙げられる。その評価には、他の医療機器にはない新たな考え方が必要となる。この場合、性能変化が適切な方向へ向かっているか、あるいは新たなリスクを生じていないか等の確認が必要と考えられる。
- ・ 今年度の目標は評価指標案を作成することであり、予備日を含めて 6 回の会議開催を予定している。
- ・ 学会連携委員の皆様は、各自学会に情報を持ち帰り、学会での意見を取りまとめる方向でご参加いただきたい。

5-2. 話題提供（講演）

(1)「医療機器の承認審査について」（PMDA）

PMDA 医療機器審査第一部の担当者から、承認審査の基本的考え方が紹介された。講演内容とその後の質疑応答の主な内容は以下のとおりである。

『講演概要』

- ・ 医療機器の定義は、薬機法第 2 条による。
- ・ 薬機法の改正によって、単体プログラムは医療機器の範疇にはいるようになった。
- ・ 製造販売承認申請においては、承認を受けた後に改良や変更が生じた場合、軽微なものを除き、一変申請し承認を受けなければならない。性能が継続的に変化する AI の場合、本規定の中でどのように取り扱うかについては、当 WG の議論の一つになると思われる。
- ・ AI のように継続的に性能が変化し、特定ポイントのみで評価することが難しい製品も、使用成績評価の制度を利用することができるかもしれない。
- ・ 承認審査では、有効性と安全性のリスクベネフィットバランスを評価し、使用目的または効果、使用方法、使用上の注意の妥当性が判断される。
- ・ 審査対象機器の臨床上の位置づけによって、示すべき有効性、安全性のポイントが変わる。
- ・ 承認審査において、有効性は臨床上のアウトカムに対する影響として評価している。性能の向上と有効性の向上は同義ではない。

『質疑応答』

- ・ 認証審査における第三者認証機関とはどのようなものか？ FDA とクラスの考え方は同じではないのか？（WG 委員）→ アメリカと日本のクラス分けの考え方は違う。日本の場合、第三者認証機関で認証される医療機器は、既に複数の同等の医療機器が承認済みで、認証基準がある製品群に限定される。Computer-Aided Detection/Diagnosis (CAD) は、基本的にはクラス 2 に該当し、それらに適用可能な認証基準「核医学装置ワークステーション用プログラム等基準」が存在するが、認証基準の対象範囲を超える機器の場合は PMDA で審査している（PMDA）。

(2)「人工知能を利用した医療機器の承認審査における課題」(厚労省)

厚労省の担当者から、AI を利用した医療機器の承認審査における課題について紹介された。主な講演内容は以下のとおりである

- AI 活用については、官邸、内閣府、総務省、経産省、文科省で会議が立ち上がっており、期待が高まっている。
- 厚労省でも「保健医療分野における AI 活用推進懇談会」を開催している。
- AI 技術を用いた医療機器への対応については、厚労省としての目標の中に組み込まれており、この WG の成果は非常に重要である。なお、医療機器への該当性については、監視・麻薬対策課が対応することになる。
- 医療機器プログラムについては、近年の法改正により、薬機法の規制対象として取り扱われるようになった。
- 個々のプログラムの医療機器への該当性は、治療方針等の決定への寄与とリスクの蓋然性の 2 軸で考えたときの位置づけによって判断される。
- 医療機器プログラムの特性として、医師が直接プログラムにアクセスする場合は、医療機器プログラムの提供（使用権の提供）となり、医療機器として規制の対象となるが、医師の依頼により企業がプログラムを操作する場合は、解析サービスの提供となり、薬事法の対象外となる。
- 「保健医療分野における AI 活用推進懇談会」では、AI による診療支援と医師の判断との関係性の整理を行っている。例えば、1)診断確定や治療方針の最終的な意思決定は医師が行い、責任も医師が負うべきこと、2)その前提としては医師に対して適切な教育を行い、安全性を確保することが必要であること、3)保健医療分野における AI 開発には医師の関与が必要と思われること等が議論された。
- 昨年承認された疾病診断用プログラム「ハートフロー FFR_{CT}」では、心臓の CT データを Web 経由でハートフロー社に登録・転送し、解析を依頼すると、ハートフロー社に存在するプログラムを使用してシミュレーション解析が行われ、解析結果が Web 経由で返送される。医師は解析結果を参考として治療方針を決定するという流れとなる。プログラム使用権の提供は医療機器プログラムの販売とみなされ、医療機器に該当すると判断された。
- ハートフロー FFR_{CT} では、「医療機関から転送された CT データ等の目的外の使用を妨げる手続き・適切な管理を行うこと。最新のセキュリティ対策を講じること。」等の承認条件がつけられている。
- 医療機器プログラムの申請に関するガイダンス（平成 28 年 3 月 31 日医療機器再生医療等製品担当参事官室事務連絡）に審査上の論点が見られている。また、「コンピュータ診断支援装置に関する評価指標」も発出されているので、参考にしていきたい。
- AI 特有の検討事項としては、性能の評価、性能変化の検証、性能変化の管理、問題発生時の対応、性能の変化による使用目的の承認からの逸脱、使用者への注意喚起、トレーニング等が考えられる。

(3)「人工知能は人間を超えるか」(招聘講演)

東京大学・松尾豊先生から、人工知能の現状等が紹介された。主な講演内容は以下のとおりである。

- DL は、1)認識、2)運動の習熟、3)言葉の意味理解の順で実用化されていくと考えられる。
- DL は画像認識が可能になったところにブレイクスルーがある。従来の人工知能で

は、人間が特徴量を抽出してモデル構築していたが、DL では、データに基づいて特徴量を自動的に獲得するようになった。(2012,Google の猫)

- ・ 2012 年以降、DL の導入により、認識のエラー率は格段に下がっている。2015 年 2 月には人間の精度を超えたと言われている。(ヒト 5.1% vs DL 3.1%)
- ・ 運動の習熟については、DL と強化学習の組み合わせで行われている。試行錯誤の過程で、より良い成果を出せば、報酬を受け取る作業を繰り返して、精度を向上させていく。
- ・ 今後、言葉の意味を理解した言語処理ができるようになると思われる。
- ・ 従来、ヒトの眼による識別作業が必要であった産業分野（農業、建設業、食品加工業等）においても、DL を利用した画像認識技術と機械系やロボット系技術を組み合わせることで自動化が可能になると考えられる。しかしながら、これらに対応できるハードウェアは開発途上である。それに対し、画像処理のみで目的が達成される医療画像については、マーケットが大きいこともあり、ベンチャーが多く参入してきている。今後、様々な製品が開発されると予測されるため、対応を急ぐ必要がある。
- ・ 多量の学習が必要であることから、学習自体はサーバー側で行われると思っている。
- ・ オンライン学習に伴うリアルタイムでのバージョンアップは難しいと思っている。一回の学習にも相当の時間を要する上、ハイパーパラメータを含めてチューニング数が多いことから、バージョンアップは数ヶ月に一回程度になると予想している。
- ・ 必要な学習データ数に関しては、教師あり学習の場合、一般的にカテゴリーごとに約 5,000 のラベル付き事例で許容できる性能を達成すると言われている。しかしながら、現時点で人間の性能に匹敵するもしくは超える性能を得るためには 1,000 万枚の画像データ＋アノテーションが必要であると考えられている。
- ・ 保健医療分野において、診断の自動化は難しいと考えている。しかしながら、画像認識の精度の向上に伴い、画像診断自体の実用化は早いと思われる。
- ・ 本 WG での取り組みは、DL 活用の最も早い先行事例となる。様々な産業に大きな影響を与えることから、とても重要なものであると考える。

5-3. 総合討論

(1) 現時点で想定される問題点や疑問点及び今後の検討課題

AI を利用した画像診断技術に関する現時点で想定される問題点や疑問点及び今後の課題について自由討論した。主な意見、討議内容は以下のとおりである。

- ・ 現行の CAD は、コンピュータで定量的に解析された結果を『第二の意見』として利用する『医師による診断の支援』として位置付けられているが、昨年 12 月に開催された北米放射線学会において、医師が大量の画像を見るのではなく、データのふりわけが行われ、医師はコンピュータによって示された画像のみを見ればよいというタイプの機器が提案されていた。医師が見る画像が制限される機器もターゲットとして考えていく必要があるが、このときの責任の所在はどこか？(WG 委員)→厚労省の「保健医療分野における AI 活用推進懇談会資料」に示されている通り、医師が決定し、責任を持つと理解してよいと考えられる (WG 委員)。
- ・ 多くが健常者である場合のスクリーニング用と、病態を有する患者の診断用とでは同じ画像診断でも考え方が異なると思うが、全体を包括する評価指標を作るのか？スクリーニング用の指標と高度な診断用の指標は分けて考えないと、意見が出にくい (WG 委員)。
- ・ 評価指標案が長期に渡って使用されるようであれば、近年の技術動向を踏まえた内容にすべきである (WG 委員)。

- ・ 医師が意思決定を行うことと、すべての画像を見ることは意味合いが違う（WG 委員）。
- ・ 細胞診システムでは、ベクトンディッキンソンから **Focal Point** というクラス 1 機器が上市されている。この機器は自動スクリーニング装置である。原則、スクリーナーが再度確認することになっているが、現状では **Focal Point** で異常がないと判定された症例に関しては再確認されていない。また、「細胞くん」という AI も精度がよく、陰性と判定された症例の病理を見ることはない。一方、病態を有する患者の細胞診でも同様の処理がなされている実状があるので、そのことも勘案して評価指標を作成して欲しい（WG 委員）。
- ・ 胸部単純 X 線写真による正常の判定については、統計学的有意差をもって正常と判定できるレベルに到達していることが、NIH の研究で証明されていることも踏まえて欲しい（WG 委員）。
- ・ 病理分野においては、科学知見の進歩とともに、診断名や疾患カテゴリーが変更された事例があり、今後も同様のことが起こることが予測される（WG 委員）。→変更前の情報で教育され、承認された機器の承認取り消しのタイミングについても討議するべきである（WG 委員）。→診断名の変更に伴い、新たな診断基準が確立された場合、その旨を関係諸方面へ一斉に情報提供する必要がある。また、その基準を満たしていなければ、機器の承認が取り消される流れになる可能性が考えられる（厚労省）。
- ・ データラベル自体の張替は難しくないが、細分化したラベルは張替が難しいことを留意する必要がある（オブザーバー）。
- ・ AI の学習に使用する画像データの所有権は誰に帰属するべきか？（WG 委員）→医療情報は公共性が高いと考えられることから、国が一括管理するか、ある条件を満たした場合は、不特定のユーザが使用できるような、なんらかの仕組みを作る必要があると思われる。海外企業は装置を安価に提供する代わりにデータを収集する戦略を取る可能性がある（オブザーバー）。
- ・ 患者から提供されたデータが、目的外のことに転用される懸念がある（WG 委員）。
- ・ 個人情報保護法により、データを収集することに対する倫理的な問題がある（WG 委員）。→厚労省の個人情報保護法担当者によれば、過度に個人情報として扱って、研究の推進を妨げるようなことはしたくないとのことであった。この場で議論しても結論は出せないで、メーリングリスト等で評価指標案に入れ込む個人情報に関する確認事項を決めるのがよいのではないかと（WG 委員）。→個人情報に関しては、この WG で決められる事項ではない。診療情報に関する取り扱い責任は医師や医療機関にあるが、一旦システム改良に使うと責任者が異なることになる。特に、企業が情報取り扱い主体になれば、法的な取り扱いも変わってくる。情報の扱いについては、個人情報保護法の解釈も含め、ガイドラインに指針的な事項を書き込む必要があると思われる（WG 委員）。

(2) 評価指標案作成のために討議すべき点

事務局より資料 7 を用いて、本 WG で評価指標（案）を作成するにあたり討議すべきと考えられる以下の事項について説明し、6 月末を目処に WG 委員の意見を求めることとした。

- ・ 対象となる機器の範囲
- ・ 勘案すべき項目
- ・ 変化する性能の評価方法
- ・ AI に特化した新しいリスクとその防止法
- ・ 取り扱うデータ

- ・ 参考となるガイドラインの有無

5-4. 開発 WG の活動計画

開発側事務局を担当する坂無氏から、開発 WG の状況が報告された。主な内容は以下のとおりである。

- ・ 審査 WG と平行して開発 WG を開催する。
- ・ 開発 WG では、これまでに Computer-Aided Detection (CADE) 及び Computer-Aided Diagnosis (CADx) のガイドラインを作成した。
- ・ 開発 WG では、審査 WG から 3 名程度兼担していただくことを予定している。
- ・ 開発 WG には企業委員も参加可能であるため、JIRA から招聘する予定である。
- ・ 審査 WG と開発 WG の合同会議を設定する予定である（一回目は、審査 WG 第 3 回会議開催日である 9 月 28 日とする予定）。

5-5. その他

- ・ すでに医療機器として承認された機器をリストアップしてほしい。CADE の乳がん診断装置は上市されているが、CADx に相当する製品は存在しないと思われる (WG 委員)。→公開されていない情報は難しい (PMDA)。→WG 委員の方々と、該当するクラス 1 製品や、機器の附帯機能として上市されているソフトウェアをご存知であれば情報を提供してほしい。公開情報等を含め、得られた情報の整理を事務局と PMDA で行う (事務局)。
- ・ 第 2 回会議 (8/25) では、事務局から評価指標案たたき台を提示する予定である。第 2 回会議で討議した後、学会へ持ち帰り検討していただきたい (事務局)。

5-6. 今後の会議日程について

第 2 回会議： 8 月 25 日 (金) 14:00 – 17:00

場所：TKP 品川カンファレンスセンターANNEX (予定)

以上

次世代医療機器・再生医療等製品評価指標作成事業
人工知能分野 審査 WG
平成 29 年度 第 2 回会議 議事概要

1. 開催日時：平成 29 年 8 月 28 日（金）14:00 -17:11
2. 開催場所：TKP 品川カンファレンスセンターANNEX カンファレンスルーム 4
（東京都港区高輪 3-13-1 TAKANAWA COURT 3F）
3. 出席者（委員：座長以下五十音順・敬称略）
座長：橋爪 誠（九州大学）
委員：佐々木 毅（東京大学）、清水 昭伸（東京農工大学）、田中 聖人（京都第二赤十字病院）、中田 典生（東京慈恵会医科大学）、藤田 広志（岐阜大学）、待鳥 詔洋（国立国際医療研究センター国府台病院）、横井 英人（香川大学）
厚生労働省（厚労省）：中井 清人、青柳 ゆみ子
医薬品医療機器総合機構（PMDA）：高江 慎一、菅原 明彦、加藤 健太郎、岡崎 譲、
今川 邦樹、江原 輝喜、若尾 りか
オブザーバー：坂無 英徳（産総研）
事務局（国立医薬品食品衛生研究所）：畠島 由二、中岡 竜介、加藤 玲子
4. 配布資料
資料 1：第 2 回審査 WG 会議議事次第（案）
資料 2：委員名簿
資料 3：第 1 回審査 WG 会議議事概要（案）
資料 4：人工知能を利用した画像診断機器に関する評価指標案作成にあたっての課題
資料 5：人工知能を利用した画像診断機器に関する評価指標案作成にあたっての方向性整理（事務局案）
参考資料：ヒヤリング対象企業からの配布資料
5. 議事内容
 - 5-1. 開催挨拶及び資料確認
座長及び事務局挨拶後、事務局より配布資料の確認・説明及び初参加者が紹介された。
 - 5-2. 企業ヒヤリング
AI 技術を利用した製品開発状況を把握するため、関連企業からヒヤリングを実施した。
 - 1) エルピクセル
 - 2) 東芝メディカルシステムズ
 - 3) 島津製作所
 - 4) 日立製作所
 - 5) 富士フイルム
 - 6) オリンパス

5-2. 総合討論

資料 5「人工知能を利用した画像診断機器に関する評価指標案作成にあたっての方向性整理（事務局案）」に基づいて自由討論した。主な討議内容は以下の通りである。

(1) 本評価指標の対象

- ・ 企業ヒヤリングより、必ずしも人工知能がクラウド上に存在している必要はないと思われる。
- ・ 企業ヒヤリングでは、単なる画像処理、または治療を目的とした画像処理に AI を利用している等、使用目的が診断支援に該当しない機器も紹介された。これらの機器は今回作成する評価指標案の対象となるのか？ 対象範囲を明確にするために、診断支援の定義が必要と思われる。（委員）→ 画像診断支援システムの定義は「コンピュータ診断支援装置に関する評価指標」に記載されている。AI 使用の有無に拘らず、その定義を利用すれば良い。（委員）→ 画像処理のみを行う製品の取扱いを明確にする必要がある。（委員）
- ・ 画像検索は対象に含めるか？（委員）→ 企業ヒヤリングから判断した場合、対象に含めた方が良いと思われる。（委員）→ 画像に限らず、医療情報を利用して診断結果を導く製品は、排除するという理解で良いか？（委員）
- ・ CAD への入力用画像の前処理に AI を利用している既存製品も対象とするのか？（委員）→ 従来の CAD でもニューラルネットワークを利用している製品が存在することから、やはり AI の定義が必要ではないか？（委員）→ 従来のニューラルネットワークとディープラーニングの両方を対象とするか否かについては未検討である。（委員）→ 今回は、新たに AI を使うにあたっての考え方を整理してまとめる方向で議論を進めれば良い。（委員）
- ・ 画像処理のみに AI を使っている製品等、事務局提案の 3 つの機能の範疇に入らないものは除外になるか？（委員）→ 企業ヒヤリングでは、治療支援技術と考えられる製品も存在した。今回の対象は診断支援機器であり、治療支援や手術支援等は対象外と考える。（委員）→ 内視鏡画像による診断は微妙な立ち位置になる。（委員）→ 今回提案された対象範囲に内視鏡も含まれると判断する。（委員）
- ・ ここまでの討議を受けて、事務局案から「クラウド上に存在する」の文言を削除し、下記の表現に変更することとした。

『本評価指標は、人工知能を利用して、

- 疾患名まで特定せず、病変の疑いがある部位のみ検出する
- 病変及び疾患候補の絞り込みを行う
- 疾患候補の絞り込みに加え、それらの重み付けまで行う

医用画像診断支援システムを対象とする。』

(2) 評価に対して留意すべき事項

- ・ 今後新たに上市される製品にも対応できるようにするため、具体的に何を評価すれば良いか分かり易く「基本的考え方」を記載しておくが良い。（委員）
- ・ 病理画像も含めておいた方が良いと思われる。（委員）
- ・ 学習用データの種別や解像度等については、現在、企業等毎に独自に開発、または設定している事例が多いため、早急に基準を示す必要がある。（委員）→ 必ずしも解像度が高い方が優れているとは限らないことから、使用した学習データの条件を提示することは重要と思われる。（委員）
- ・ 用途（first reader, second reader, concurrent）及び目的（スクリーニング、診断支援）に応じて、評価方法も相違すると思われる。モダリティによっても評価方法

が異なることがあると考えられることから、本 WG 内で「基本的考え方」について議論し、評価指標案に記載する文章を作成していただきたい。(事務局)

- ・ 共通部分と機器毎の目的に応じた部分に分けて記載することが重要と思われる。(委員)
- ・ 企業ヒヤリングで紹介された製品は、年に 1,2 回のバージョンアップを念頭においていた。リアルタイムで性能が変化する製品の上市は現時点で想定されていないと思われたため、これを踏まえて整理すると良いのでは？(事務局)
- ・ 性能が変化する場合でも、ある程度決まった範囲内での変更であれば届出で良いと思われるが、範囲が決められない場合は、一部変更届が必要なのか、PMDA 等への相談を経て届出で済むのか、行政的な判断となる。ただし、対象疾患が変わる等、機器の本質が変化した製品は一変が必要である。(PMDA)
- ・ 単なるアルゴリズム変更の場合は、それを検証できないため、性能が前バージョンと比較して劣らないことを検証結果として示すことになる。→ 前バージョンの製品を評価した時のテストデータを使用して、改良バージョンの性能を検証し、非劣性を示せば良いのか？(委員) → 同じテストデータを 2 回使用することになり、問題ではないか？(PMDA) → 学習用データではなくテスト用データなので良いのでは？(委員) → この問題を懸念している企業も存在するため、今回議論の対象としてほしい。(PMDA) → 診断薬の場合は、既存製品との違いを鑑みて同等以上の性能を有しているか確認する。既存データを使用しないと比較できないため、やはり同じテストデータを用いて一度検証するしかない。(委員) → 改善ではなく、非劣性が証明されれば良いが、非劣性が出ないようなデータを恣意的に用いて性能や精度を比較することができないようにして欲しい。(PMDA)
- ・ 使用する検証用データの種類の種類は非常に重要である。共通に使用できるデータベースが存在していれば良いが、企業が独自に収集したデータを用いる場合、大企業と中小企業ではデータの数や質に差異を生じることが懸念される。(委員)
- ・ 検証にあたり、データ数は多い方が良いが、データの偏りがないように工夫することが必要と思われる。例えば、10,000 のデータで検証する際、10,000 データで一回試験を行った場合と、任意に 1,000 データを抽出し、10 回試験を行った時の結果に差がなければ、性能・精度ともにデータの偏りもなかったと考えられる。(PMDA)
- ・ 検証に必要なデータ数については、具体的な数字ではなく、それぞれのモダリティや手法毎に適切な数を提示するための考え方を記載していただきたい。(PMDA)
- ・ 画像データ数のみでなく、病変、または正常に関する画像の質も重要になる。(委員)
- ・ 事務局から、資料 5「(2)/e 項」に「性能検証用のテストデータは、対象とする母集団の特徴を十分にカバーする質と数のデータが必要だが、機械学習の方法や対象疾患に応じてその必要数は異なり、一律に定義することが困難であるため、使用したデータセット及び検証方法を記載すると共に、その妥当性検証スキームを個別に記載すること』を要求する評価項目を設定している旨が紹介された。この項目の記載内容が、これまでの議論に対する回答となることが提案・了承された。

会議終了予定時刻となったため、討議を打ち切り、その他の項目を含めた資料 5 全体について、各分野の専門的な立場からコメントを提出していただくこととなった。

5-3. 次回会議までの流れ

- ・ 資料 5 に各委員から現在までに寄せられたコメントを簡潔にまとめてある。内容を再確認し、1 週間を目処にメーリングリストを利用してコメントを送付する。→ 放

射線、病理、消化器内視鏡では、使用するモダリティも異なり、それぞれ独自の考え方がありと思われることから、専門的見地からの意見が欲しい。(座長、事務局)

- 評価指標案の構成としては、共通の「基本的考え方」の下に各分野に特徴的な要件等を記載することを想定している。(事務局)
- 新たに追加されたコメントを事務局側で整理し、メーリングリストを使用して WG 全体へ配信する。
- 収集した情報に基づいて、事務局側で評価指標案叩き台を作成し、次回会議で提示する。

5-4. 審査 WG 第 3 回会議（第 1 回審査/開発 WG 合同会議）について

日時：9 月 28 日（木） 14:00 – 17:00

場所：AP 東京八重洲 会議室 K（予定）

以上

次世代医療機器・再生医療等製品評価指標作成事業
医療機器等に関する開発ガイドライン（手引き）策定事業
平成 29 年度 人工知能分野 審査・開発 WG 第 1 回合同会議
（平成 29 年度 審査 WG 第 3 回会議）議事概要

1. 開催日時：平成 29 年 9 月 28 日（木）14:00 -17:45
2. 開催場所：AP 東京八重洲通り 11 階 L+M 会議室
（東京都中央区京橋 1-10-7 KPP 八重洲ビル）
3. 出席者（委員：座長以下五十音順・敬称略）
<審査 WG>
座長：橋爪 誠（九州大学）
委員：佐々木 毅（東京大学）、清水 昭伸（東京農工大学）、田中 聖人（京都第二赤十字病院）、中田 典生（東京慈恵会医科大学）、藤田 広志（岐阜大学）、待鳥 詔洋（国立国際医療研究センター国府台病院）、横井 英人（香川大学）
厚生労働省（厚労省）：青柳 ゆみ子、他同行者 3 名
内閣官房 健康・医療戦略室：岡本 利久、他同行者 2 名
医薬品医療機器総合機構（PMDA）：岡崎 譲、加藤 健太郎、古森 亜矢、谷城 博幸、
宮崎 生子、今川 邦樹、江原 輝喜、若尾 りか
事務局（国立医薬品食品衛生研究所）：畠島 由二、中岡 竜介、加藤 玲子
<開発 WG>
座長：横井 英人（香川大学/審査 WG 委員も兼任）
委員：片倉 由紀子（富士フイルム株式会社）、椎名 毅（京都大学）、清水 昭伸（東京農工大学/審査 WG 委員も兼任）、縄野 繁（国際医療福祉大学）、野波 徹緒（オリンパス株式会社）、藤田 広志（岐阜大学/審査 WG 委員も兼任）、古川 浩（東芝メディカルシステムズ株式会社）、宮崎 靖（株式会社日立製作所）、森 健策（名古屋大学）、諸岡 直樹（株式会社島津製作所）
経済産業省：中島 教順
オブザーバー：浅沼 直樹（AMED）、宮本 大誠（PMDA）、田中 弘（日本画像医療システム工業会）、吉田 眞司（オリンパス株式会社）
事務局（産業技術総合研究所）：鎮西 清行、坂無 英徳
4. 配布資料
<合同会議配布資料>
資料 1：議事次第（案）
資料 2：出席者名簿（審査 WG）
資料 3：出席者名簿（開発 WG）
資料 4：講演資料 1「個人情報保護法の概要」（厚生労働省）
資料 5：講演資料 2「医療分野の研究開発に資するための特許加工医療情報に関する法律について」（内閣官房 健康・医療戦略室）
資料 6：人工知能を利用した画像診断機器に関する評価指標案たたき台（事務局案）

<審査 WG 会議配布資料>
資料-審査 1：第 2 回審査 WG 会議議事概要（案）

<開発 WG 会議配布資料>

資料 1：第 1 回 開発 WG 会議 議事概要（案）

資料 2：（仮題）診断支援技術や医療用 AI に対する期待

資料 3：（仮題）内視鏡診断支援システムに関する研究紹介

資料 4：今年度の活動に関する事務局案

5. 議事内容

5-1. 開催挨拶及び資料確認

開会にあたり、審査 WG 座長、開発 WG 座長及び各事務局挨拶があった。事務局より本合同会議の議長として橋爪先生を推薦し、異議なく全会一致で承認された。また、審査 WG 事務局が進行役を務めることとなり、配布資料の確認及び参加者が紹介された。

5-2. 話題提供

1) 個人情報保護法の概要説明

合同会議用資料 4 に従って、厚労省より、個人情報保護法の概要と改正の主なポイントが説明された。

- ・ 個人情報保護法の概要
 - 個人の権利・利益の保護と個人情報の有用性とバランスを図るための法律である。
 - 基本理念を定めるほか、民間事業者の個人情報の取り扱いについて規定している。
- ・ 改正の主なポイント（平成 27 年 9 月改正、平成 27 年 5 月 30 日施行）
 - 個人情報保護委員会が新設され、個人情報取扱事業者に対する監督権限を各分野の主務大臣から委員会に一元化された。
 - 個人情報の定義が明確化された。→ 制定時には想定されていなかったパーソナルデータの利活用が可能となり、個人情報に該当するのかどうかの判断が困難な「グレーゾーン」が拡大したことから、個人情報の定義を明確にする必要が出てきた。
 - 個人情報の有用性を確保（利活用）するための整備により、匿名加工情報の利活用の規定が新設された。→ ビックデータ（その情報が個人につながる状態で価値を有する場合）を利活用する必要性が出てきたため、匿名化することにより利活用を可能にできる対応がなされた。
- ・ 個人情報の定義（個人識別符号）
 - 「個人識別符号」以下①②のいずれかに該当するものであり、政令・規則で個別に指定されている。
 - ①身体の一部の特徴を電子計算機のために変換した符号
⇒DNA、顔、虹彩、声紋、歩行の態様、手指の静脈、指紋・掌紋
 - サービス利用や書類において対象者ごとに割り振られる符号
⇒公的な番号（旅券番号、基礎年金番号、免許証番号、住民票コード、マイナンバー、各種保険証等）
- ・ 事業者が遵守すべき主なルール
 - これらはいくまでも利用する際のルールであり、どういう目的に使用してはいけない、もしくは使用できるということを定めているルールではない。
 - 個人情報を取得・利用する時のルール
⇒個人情報を取得した場合は、その利用目的を本人に通知、又は公表すること。
 - 個人情報を保管する時のルール

⇒情報の漏えい等が生じないように安全に管理すること。

➤ 個人情報情報を他人に渡す時のルール

⇒個人情報情報を本人以外の第三者に渡すときは、原則として、あらかじめ本人の同意を得ること。

- 本 WG では、カルテ情報等を画像情報と結び付ける際に、個人情報に該当し得るものが含まれてしまうため、それらの情報を使えなくなるのではないかという懸念が示されているかと思うが、個人情報保護法は個人情報を使用してつくった成果物、あるいはそれによって得られる利益について、何かを定めているものではない。まず、AI 機器に利用する情報を明確にした上で個人情報となりうる範囲を考慮しつつ、評価指標を作成してもらいたい。

<質疑>

- 例えば、患者の画像データから特徴を取り出し、システムを作成した場合、ある時点で某患者が情報提供を拒否した場合、そのシステムを再度作成し直す必要があるのか？（審査 WG 委員）→個人情報保護法において、法律に基づいて個人がすでに与えた情報を引き戻す際の規定は存在していない。あくまでも情報収集に関わる法律であって、成果物や利益に関しては論じていない。私見であるが、AI では、個人情報を一切含まない状態に加工された画像データ等を用いて学習していくことから、AI 自体は個人情報には該当しないと考えている。（厚労省）→学習データとして不特定多数の情報から作成された評価関数のようなものであれば、個人情報は含まれていないと考えている。また、法律に基づいて得られた同意の撤回に関しては法律で規制されておらず、お互いの契約の中でルール決めしていると考えられる。すべてが個人情報保護法等の法律で決まっている訳ではない中で、提供者と利用者の間でデータの取り扱い方法等のルールを明確にし、統計情報としたものは個人情報に該当しないことをわかりやすく提供者に説明する等の工夫をすることで、お互いに共通理解を持っておくことが大事だと思われる。（厚労省関係者）

2) 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

合同会議用資料 5 に従って、内閣官房 健康・医療戦略室より概要説明があった。

- 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（次世代医療基盤法）は、昨年の成長戦略の中で、医療情報を活用した基盤をつくり、これを法制化することが閣議決定された中、有識者で議論した上で、今年の 3 月に法案として提出され、4 月に内閣委員会で審議・成立し、5 月 12 日に公布された。施行については来年の 4 月もしくは 5 月を予定している。
- 策定の背景
 - 日本における医療情報の利活用として、現在はレセプトの形でインプット情報が中心になっているが、結果としてもアウトカムをもっと利活用できるようにしたい。
 - 欧米諸国では既に医療情報の利活用が進んできていることから、日本としても、医療情報の利活用を考える必要がある。→ 日本では医療機関の規模が小さいといったことや、多々ある保険者の大量のデータの蓄積をどう考えていくかを検討する必要がある。
 - 個人情報保護法が改正された中、学術研究は従来と同様に適用除外であるが、企業の利用等を念頭においた際、要配慮情報ということで配慮する部分と、匿名加工情報として利活用を進める部分が出てきた。このような新たなステージが出てきた中、医療情報の利活用をより推進していくためにどのようなことが

必要であるのかを議論してきた。

- 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は、医療分野の研究開発のために匿名加工して特定の個人が識別できないようにした情報を利活用することを推進していく法律ということで、次世代医療基盤法と呼んでいる。

- ・ 次世代医療基盤法の構成

- 1：基本方針の策定

- ⇒具体的にどう利活用を図るかについて、施行までに基本方針を閣議決定により決める予定である。

- 2：認定匿名加工医療情報作成業者（以下、「認定業者」と記す）

- ⇒医療情報は機微な情報であることから、安全・安心に収集・管理した上で、研究開発のニーズを損なわないようにしながら、且つ特定の個人が識別できないようにするには高い技能が必要になると考えられる。そのような能力に関する基準を設定し、それに適合する人（業者）を行政が認定するという仕組みを入れ込む。さらに、①認定事業者の責務として、目的外の使用禁止、漏えい防止のための安全管理措置を講じる、従業員へは罰則付きの守秘義務を課する、医療情報等の取り扱い委託は、主務大臣の認定を受けた者に対してのみ可能とすると同時に、②認定事業者の監督として、行政が監視する仕組みを入れる。

- 3：認定業者に対する医療情報の提供

- ⇒改正個人情報保護法では、要配慮個人情報、医療情報の第三者提供については原則オプトインということになっているが、本仕組みに関しては、その目的が医療分野の研究開発であること、匿名加工医療情報を使用すること、およびセキュリティについては国が認定していることから、本人に通知して拒否されなければ、医療機関等が認定事業者に医療情報を提供できることになっている（特例としてオプトアウト可能）。医療機関等から認定事業者への医療情報の提出は任意であることから、この制度がうまく利用され研究開発の促進につなげるためには、医療機関に情報を提供してもらうためのインセンティブをどう考えるかが課題になると思われる。

- 4：管轄する省庁

- ⇒内閣官房で法案では、医療分野の研究開発に関わる4つの省庁（内閣府、文科省、厚労省、経産省）で共同管轄することにしている。

- ・ 次世代医療基盤法の期待される効果例

- 治療効果や評価等に関する大規模な研究の実現（診療支援ソフト開発等）
- 医薬品市場後調査等の高度化、効率化（医薬品の安全対策としての MIDNET 等）

- ・ 改正個人情報保護法の施行後、匿名加工医療情報の提供については、医療機関内もしくは委託業者先での匿名加工後は、利活用に本人の同意なく医療情報を提供することも可能であるが、下記のような課題がある。

- 匿名加工の専門家でない医療機関が匿名加工工程における全責任を負うことになる。
- 医療機関内を通じて他科との情報をつなぎ合わせることは困難である。
- 委託業者の技能を評価する基準が定められていない。

- ・ 個人の権利利益の保護に配慮しつつ、匿名加工された医療情報を安心して円滑に利

活用することが可能な仕組みを整備した。

- 高い情報セキュリティを確保し、十分な匿名加工技術を有するなどの一定の基準を満たし、医療情報の管理や利活用のための匿名化を適正かつ確実に行うことができる者を認定する仕組み（＝認定匿名加工情報作成業者）を設ける。セキュリティや匿名加工の能力を国が認定することで、匿名加工の責任は認定業者が負う。
 - 医療機関等は、本人が提供を拒否しない場合、認定業者に対し、医療情報を提供できるとされ、認定業者は、収集情報を匿名加工し、医療分野の研究開発の用に供することが出来る。また、複数の医療機関情報の突合の上、匿名加工することが可能になる。
 - これらの匿名加工情報を企業・アカデミア・行政のそれぞれが利活用した成果が医療現場に戻ってくることにより、最適医療の提供や医薬品副作用等の早期把握による安全性の向上を患者・国民へ還元する効果が期待される。
- ・ 利活用の基本的考え方
 - 匿名加工された医療情報の円滑かつ公正な利活用を、基本方針（閣議決定）と認定基準で確保する。
 - 認定業者は情報利活用性の目的を個別に認識し、健康長寿社会の実現に資する公益性の高い研究開発に優先的に対応する。不適切な利用は排除し、意図せざる目的への転用を防止する。（患者から情報を任意で集めるため、信用を損なわない対策を講じ情報収集が円滑に行われるように努める。）
 - 情報の収集加工提供に要するコストの利活用者への転嫁を基本とし、認定業者に過度な利潤を生じさせない。付加価値をつけるのは、あくまでも利活用者。
 - ・ セキュリティについては、事前及び万が一の場合の事後の対応を徹底する。具体的には、組織・人的要因によるリスク排除、基幹システムのオープンネットワークからの分離、多層防御・安全策の導入（個人情報の暗号化や緊急時の対応・監督体制の確保を含む。）を徹底させ、且つ医療ニーズに精通し高い匿名加工技術を有する業者のみを認定することになっている。
 - ・ 本仕組みはこれまでにない新しい仕組みであることから、認定業者による匿名加工事業は、適正な規模で始め、成果等を確認しつつ徐々に拡大していく予定である。
 - ・ 個人情報保護法における匿名加工
 - 匿名加工情報とは、特定の個人を識別することができないように個人情報を加工して得られる個人に関する情報であって、該当個人情報を復元することができないようにしたもの。
 - 個人情報保護委員会が定める基準（1. 特定個人を識別可能な記述、2. 個人識別符号、3. 個人情報の連結符号、4. 特異な記述を削除する、または 1～4 を復元できない方法により他の記述等に置き換える）に沿って的確に匿名加工を行う能力を有する事業者を認定する。
 - 医療匿名加工情報も定義内容は個人情報保護法と同じである。（特定の個人を識別できないように加工して得られる情報）
 - どのような基準を満たすべきであるかについては、基本的には個人情報保護委員会が定める基準を踏襲することになると考えられるが、医療分野での適用により合致するように、記載を削除したり置き換えたりする処置が必要になると思われる。

- 「医療分野における特異な記述」については、施行（来年 5 月）までに学会等の意見も取り入れながら検討し明らかにしていく予定である。さらに、加工技法自体の進歩も考えられるので、施行後も含め、どのような情報について、どのような匿名加工を施していけば良いのか、また、それを研究開発のニーズを損なわないように、どう実現していくかを検討し、継続的に具体化していく予定である。

<質疑>

- ・ 過去に取得した情報で同意を得ていない場合や使用目的が違っても医療機関内で匿名加工すれば第三者に提供することは可能ということか？（審査 WG 委員）→ その通り。改正個人情報保護法の目的の一つにビックデータの活用があり、ルールに従った匿名加工情報の第三者提供に際しては特に同意等の必要がないというのが新しい点である。（内閣官房・厚労省関係者）
- ・ 個人情報保護法下ではオプトイン扱いのものが、次世代医療基盤法下ではオプトアウトでよいということはどういうことか。（審査 WG 委員）→ 個人情報保護法下では、要配慮個人情報である病歴等の情報の第三者への提供は、原則として本人の同意がないと提供できないことになっている（オプトイン）。しかしながら、次世代医療基盤法に基づいて医療機関から認定業者へ個人情報を提出する場合は、本人に通知をして拒否されなければ個人情報を提供できるようにしたということである。（内閣官房）
- ・ 学術目的の場合の除外規定というものはあるか？（審査 WG 委員）→ 個人情報保護法には学術という例外規定を置いてあるが、次世代医療基盤法にはない。そもそも次世代医療基盤法は、医療分野の研究開発のために匿名加工をする認定事業者に個人情報を渡す行為を容認するものであるため、大学であっても企業であっても使用目的が医療分野の研究開発にあたるのであれば、匿名加工情報を使用することができる。（内閣官房）
- ・ 学術目的で二次提供までであればオプトアウトは許容されているのか？（審査 WG 委員）→ そもそも、学術研究については個人情報保護法では第四章の規定において適用除外になっており、その取り扱いについては学術世界のルールで決まることだと認識している。（内閣官房）
- ・ 今後、学会がデータベースを預かるような法人を設立した場合、そのような団体でも、一定の基準を満たしていれば事業者認定を出してもらえるのか？（審査 WG 委員）→ 法律の中では、法人格があること、欠格要件に当てはまらないこと（個人情報関係で過去に犯罪に関与していないこと）、医療分野の研究開発に資するような匿名加工ができる能力を有すること、情報セキュリティの措置が出来ることが要件となっている。ただし、基本的には、要件基準を高く設定し、その基準を満たした限られた業者が運営していくことになる想定している。（内閣官房）
- ・ 院外の認定業者に個人情報を渡す際は匿名化しておかなくてはならないのか？（審査 WG 委員）→ 次世代医療基盤法は、病院から認定業者に対して個人情報の形のままで渡すことを認めた法律である。（内閣官房）
- ・ 拒否の機会は与えられているが、匿名化後に拒否したとしてもその情報の使用を過去に戻って使用不能とすることはできないと思うが？（審査 WG 委員）→ 必ず匿名化された情報に戻れるようにリスト化しておく必要があるのではないかと。（審査 WG 委員）→ 変換テーブルを持たないことを提示し、仮に外にデータが漏洩しても個人に紐付けできないようにしていることから安全であるという形で同意を取り、それを元に倫理審査を通して研究に使用している。（審査 WG 委員）→ 例えば、がん登録では個人情報保護法第 23 条（第三者提供）において、法律の定めがある場

合は同意が必ずしも必要はないという例外が定められている。また、連結可能匿名化は個人情報情報を抹消する操作ではないことが、昨年の研究指針作成時の議論で肝になった部分である。(厚労省関係者)→変換テーブルがあるということは連結可能ということになり、匿名化となくなるが？(審査 WG 委員)→詳細を確認する必要がある。(審査 WG 委員)

- ・ 希少疾患の画像についてはたとえ匿名加工しても、特定の個人が同定されるのではないかと懸念されるが？(審査 WG 委員)→どの程度「まれ」だと、そのような問題が生じるのか、現時点では具体的事例が明らかになっていないと認識している。匿名加工情報の定義では、特定の個人を識別できないときに、一般人をもって具体的な人物との間に同一性を認めることができるかどうかとなっている。希少疾患の程度にもばらつきがあるので困難ではあるが、「まれ」と言った時に、一般市民が特定の個人を特定できるかと言えば、普通はできないと思われる。指摘事項については、許容範囲をまず明らかにしていくことが必要だと考えている。医療分野の匿名加工情報は研究開発の分野で使用されることを念頭においており、一般市民にまで渡っていかないようにするため、情報提供対象者の範囲を検討する必要があると思っている。(内閣官房)
- ・ 海外には匿名加工情報をデータベース化している研究所もあるが、その情報はこの法律の範疇外か？(審査 WG 委員)→事業者と相手方との契約関係になると思う。フリーに使ってよいとなったらその情報は様々な用途に流用され、結果として事業者の経営基盤は安定しないと思われる。欧米の仕組みを研究する必要があると思っているが、オープンデータになればなるほど、悪意ある人間が介入する可能性を鑑みて、統計に近い情報にならざるを得ないと思われる。利用の範疇を限定し、その範疇に応じた情報の開示範囲についての議論も必要になると思っている。(内閣官房)
- ・ 現在、個人情報保護委員会が情報を出してよいと認めている海外の国は存在していないという理解でよいのか？(開発 WG 委員)→現在のところない。(個人情報保護委員会事務局)
- ・ 画像が入っている医療機器の修理の際、医療機関の委託ということで、医療機器メーカーが院外へ持ち出すことは許されていると思うが、海外メーカーへの委託も許されているのか？(開発 WG 委員)→個人情報を取り扱うか否かが判断基準の一つになってくる。しかしながら、修理にあたり個人情報も扱う際は、委託の範囲内になり法律で認められている。ただし、海外に出す場合は、それなりの基準を設ける必要はある。海外の事業者が我が国と同様の個人情報を保護できる基準をクリアしていることを証明する必要がある。個人情報保護法が遵守されている範囲においては海外への委託も可能と考えられる。(個人情報保護委員会事務局)
- ・ 学会が公開した画像データベースを大学が使用する場合と企業が使用する場合とでは、それぞれどのように考えるのか？(開発 WG 委員)→学会も学術団体と見なされる。民間企業や私立病院等も研究主体と見なすことができ、それらが学術団体との共同研究に属すると認められる場合には、学術研究目的での個人情報等の利用(質問のケースではデータベース利用)は認められる。ただし、営利事業への転用は学術研究とはみなされない。(個人情報保護委員会事務局)
- ・ 研究の目的によっては、性別や生年月日等の情報が必要になる場合があると想定される。場合によっては、生年程度は残しておいてもよいと考えて問題ないか？(審査 WG 委員)→研究開発におけるニーズを損なわないよう、適切な情報の取捨選択について検討する必要があるが、基本的には個人が特定できなければ良い。(内閣官房)
- ・ 診断支援の中に遺伝子データを含んでいる場合はどう考えればよいのか？(審査

WG 委員)→現状、診療の場でフルゲノム情報を持っていることはないと思われる。遺伝情報は個人識別情報にあたることから、個人識別情報に関わる基準をクリアしていることを前提に、遺伝情報と診断データとをつなぎ合わせてどこまで使用できるかについては議論していく必要がある。(内閣官房)

- ・ 次世代医療基盤法に基づいて匿名加工された情報の取り扱いには、さらに利活用に規制がかかることがあるのか？(開発 WG 委員)→個人情報保護法と同様に匿名加工情報について照合して再特定する行為のみが禁止されている。ただし、利活用にあたっての条件設定として、使用者に制限をかけるとか、使用者が約束に違反した場合は、以降、認定業者から情報を得ることができなくなる条件をつけるなどの工夫が必要と思われる。また、認定業者には、個人情報及び匿名加工情報の安全管理の措置の基準をかけており、さらに利活用先に渡す際の条件設定や渡った後のトレーサビリティの確保についても、その基準を設けることが必要になると考えている。(内閣官房)
- ・ 収集した画像や患者情報の所有者は誰になるのか？(審査 WG 委員)→個人情報保護法等においても基本的には所有権という概念はなく、誰が保有しているかということになる。その視点において、画像や患者情報は認定業者が保有していることになる。保有場所に拘わらず、患者の個人情報を保有者から第三者に渡す際に、使用目的に応じた患者の同意取得形式を考慮した法体系が形作られていると理解している。(内閣官房)→データをやり取りする際にかかってくる法律という理解で良いか？(審査 WG 委員)→ その通り。(内閣官房)

審査 WG 事務局注)

本議論を受けて、会議後、本省担当者から以下の意見があった。

- ・ 個人情報保護法で法令に基づく情報提供は可能とされている。次世代医療基盤法はその提供を「可能」とする法令なので、オプトアウト方式で個人情報提供が可能である。
- ・ 基盤法に基づいて提供者が「拒否」した場合は、提供者の医療情報がそれ以降「加工業者」に提供されなくなるのみである。それ以前に提供し匿名化された情報は既に個人情報ではないのでその利活用は妨げられない。
- ・ 匿名化により個人との紐付けはできなくなっており、基盤法でも管理表(変換テーブル)の作成・管理を義務付ける規定はない。むしろ、基盤法第 18 条で匿名化情報と他の情報との再照合は禁じられているため、管理表の作成は業者のリスクにしかない。
- ・ 以上のことから、WG 委員らから示された懸念は当たらない。

5-3. 総合討論

審査 WG 会議用資料 6「人工知能を利用した画像診断機器に関する評価指標案たたき台(事務局案)」に基づいて討論した。主な討議内容は以下の通りである。

- ・ 3 項 27,28 行目に「**first reader**、**second reader**、**con-current reader** が存在する」と限定的に記載されているが、今後、**interactive reader** タイプの機器が出てくると思われる。「等が存在する」と、表現を柔らかくしてはどうか。(審査 WG 委員)
- ・ 2 項 13,14 行目の記載は対象範囲を広げるため、「ここで対象とする人工知能とは、システムとの接続形態やアルゴリズムの種類にかかわらず、機械学習等により」に修正してはどうか。(審査 WG 委員)
- ・ 当初、その使用に伴い性能が変化していくことを想定していたが、人工知能の中には性能が変化しないものも含まれるとの意見があったことから、このような記載とした。(審査 WG 委員)→現在作成している評価指標案は、性能が変化する機器が

対象であり、性能変化を伴わない製品は既存通知である「コンピュータ診断支援装置に関する評価指標」を参照する方針だったのではないか？（審査 WG 委員）→基本的には CAD 通知の上乗せで、特に人工知能を利用した場合に注意すべきポイントを今回まとめることが、本 WG のミッションだと理解しているがそれでよいか？

（審査 WG 委員）→それで良い。ただし、平成 23 年に発出された CAD 通知では人工知能という言葉は使用されていないことから、例えば、人工知能の性能が変化しないタイプの製品は既存の CAD 通知で評価して差し支えないとの記載にすれば、今回は上乗せ部分の評価指標案という方針で作成できる。そのような方向性で良いか全委員に確認しておきたい。（審査 WG 事務局）

- 性能変化とは、ユーザーサイドでの使用中に自己学習して変化するとことを意図しているのか、それとも、企業がリリースした後は変化せず、次のバージョンアップ等によって変化することを指しているのか？（開発 WG 委員）→少なくともヒヤリングを行った企業においては、市販後、自律的に自己学習していく製品の開発を想定していないように思えた。ただし、海外では自己学習して、性能が変化する製品も既に取り入れられてきていることから、それも踏まえた考え方を記した方がよいのではないかという意見もあり、両者を包含することを想定して記載した方がよいと思われる。（審査 WG 委員）
- 既存機器でもニューラルネットワークを使用した製品が存在していることから、「学習する機能」の定義も記載しておく必要があると思われる。（審査 WG 委員）→ 今回の人工知能の定義の中に、ニューラルネットワークを使った製品に関する考え方を追記する必要性の可否についても討議して欲しい。（審査 WG 事務局）
- 基本的考え方については、市販後の性能担保の方法、あるいは更新する際に必要な条件等に関する情報を記載すればよいのではないか。（審査 WG 委員）
- 医療機器として承認された製品を自施設内で学習させて性能が変化する場合、医療機器として承認を与える是非について問題が生じるとと思われる。例えば、自施設内で CNN を使って学習させ、性能が良くなったと自己判断した製品を診療に活用するのは適切ではないと思う。メーカーがリリースする際、適切にブラッシュアップして、申請し直すことにした方がよいと思う。（審査 WG 委員）→病院ごとに性能が異なることは医療機器として好ましくないのではないかと意見があったことから、事務局案として、「性能変化は、安全性及び品質を担保するため、対象システムに対して一律に実施しなければならないことに加え、性能変化の内容に応じた手続きを行うことが求められる」と提案している。さらに「届出」、「一変」と考えられる具体案を、前回の議論を踏まえて追記している。この WG 内では、将来、本当にリアルタイムに自己学習して性能が変化する機器が開発された時、一律に性能を担保するための方法論や許容範囲の考え方に関して議論していただきたい。そのため、5/(3)/2項に「①定期的なアップデートによる段階的な性能変化を意図する場合」と「②臨床使用と並行してリアルタイムで性能変化を意図する場合」の留意点について整理してある。（審査 WG 事務局）
- デフォルトの性能と併存して結果を出力できるようにするのが望ましい等の記載をしておくべきではないか。ただし、学習を間違えてしまった際は、学習前（デフォルト）にリセットできることを担保する等の条件を記載しておけば、出荷後の自己学習に対応できると思われる。（審査 WG 委員）→内視鏡のように、基本的にリアルタイムで診断する機器の場合には、デフォルトと学習後の 2 種類の結果を提示することは不可能に近い。（開発 WG 委員）→内視鏡の場合、リアルタイムのソフトウェア更新は難しい現状であるが、そのような製品が開発された場合の対応法も念頭に置いて記載できるか検討して欲しい。（審査 WG 委員）→リアルタイムで性能変化する場合、性能監査の間隔に関する問題があるが、たたき台では定期的な性能

監査を要求するよう記載している。実際にリアルタイムに性能が変化する装置について、定期的な性能監査を行う場合、クラウドシステム以外の人工知能では対応できないと思われる。(審査 WG 事務局)

- リアルタイムで性能変化する場合のリスクマネジメント対策として、人工知能が誤った判断をする、あるいは使用者の誤使用の問題が生じた場合、メーカーと規制当局へ通知するような機能を必要に応じて付与するのはどうか。(審査 WG 事務局)
- 「リアルタイムで、かつ販売された対象システムが全て同一に変化」させることは実現が難しい。クラウドで、かつリアルタイムのシステムは、非常に早いデータ処理スピードが要求されると考えられることや、病院の要望に左右され、結局デフォルトのままにして欲しいということもあることから、「管理できる範囲で一律に変化」という記載があると厳しいと思う。(開発 WG 委員)
- 病院内で性能変化が起こる場合、病院ごとに患者の集団が異なるため、病院ごとに異なる最適化が起こると思われる。その一方で、単一の病院のデータで学習させていくことから、そのデータが標準化されているか否か、データのバリエーションについて企業は責任を負えない。結果として、医者の裁量の範囲内の使い方に依存していくように思われる。最終的にそれをふまえて診断するのは医者の責任であるが、医師の裁量範囲に関する観点は必要になると思われる。(審査 WG 委員) → 企業としては、ラベリングする範囲、確認方法の指定、QC データの提供、最低限の検出性能の保証、あるいは特定条件下で正解を導くことができない状況に陥れば使用を中止などの前提を付した上で、ユーザの要望に応えるシステムの開発を進めることを想定していた。(開発 WG 委員) → 機器の性能は医師に応じて好みが異なることがある。パラメータを選べる CAD もアメリカでは認可されているので、それと似ていると思われる。ただ、医師によって「性能」の意味が異なることもあるので、表現には留意を要すると思われる。(審査 WG 委員)
- 薬事規制として、一つの承認案件で性能の違う装置が医療機関の中に存在する状況が起こりえてもよいか考えることがキープポイントになる。一つの承認案件であることから、いずれの機器も同じような性能を有する必要があると考える場合、バッチ式、リアルタイムを問わず、一律にアップデートしていかなければいけないことになる。つまりリアルタイムに性能を個別に向上させていくシステムの場合、薬事的な規制が難しいことになる。これらの状況を鑑みると、リアルタイムの場合はクラウドを使わざるを得ないと思われる。この件については、薬事規制上、許容できる範囲を明確化しないとステップアップできないと考えられる。(審査 WG 事務局) → リアルタイムで性能が変化する製品が数年後に上市される可能性もあることから、あまり強い規制をかけるのは良くないと思われる。ただし、安全性等の問題がないことを示すことは必要だと思う。(審査 WG 委員) → 薬機法上、医薬品・医療機器ともにある一定の幅をもって安全性、性能を担保していると理解している。この観点で言えば、人工知能で新たに性能が変化する製品も、ある枠組みの中で規定された範囲内で変化することであれば問題ないと思われるが？(審査 WG 委員) → その幅が、1 品目の中で性能が全く違ってしまう場合まで含んでしまうと判断される場合は、別品目としての開発を企業の方をお願いしている。今回、1 品目として開発できる性能変化の幅を考えることになると思うが、現時点で PMDA は基準となる考えを持っている訳ではない。(PMDA) → 製造販売申請時の項目として、本評価指標案の中に「製造販売承認後に生じうる性能変化範囲の規定」に係る項目を挙げているが、現状のように箇条書きのみでよいのか。その範囲といってもかなり難しいと思われる。(審査 WG 委員)
- AI の場合、性能向上とは何をもって上がったと判断するか不明瞭である。例えば成功例があったとしても、気づかないところで不具合が起こる可能性もある。その

バリエーションの上限と下限の評価が難しいと思われる。常に同じ結果が得られるベンチマークなどで性能確認する方法を確立できればよいが、そもそもベンチマークを提供することが難しいと思われる。最終的には医師が責任を持つとは言え、製品である以上、製造者が責任を全く負わないことはないと思う。自由度を高めて開発を進めるのは、評価法が確立しない限り難しいのではないかと危惧している。(開発 WG 委員) →ある程度の幅の中で診断支援機器が規定されるのは仕方がなく、薬機法においても幅を持っていたとしても特に問題はないのではないかとと思う。(審査 WG 委員)

- ・ 人工知能は進歩が速いので、今後様々な幅の製品が上市されると思われる。明文化することは難しいが、より進化した AI が出現した際にも、医療機器として取り扱えるかどうかを本評価指標案の中に上手く記載すべきだと思う。(審査 WG 委員)
- ・ 自己学習に対して規制をかけるのであれば、AI が新たなデータを取り入れて賢くなることを止めなければならない。海外から、学習するほど賢くなるという宣伝で販売される製品が上市された場合、今回厳しい規制をかけてしまうと、そのような製品が日本で使用できないことになりかねないことから、やはり性能がリアルタイムで変化する製品についての判断基準を本評価指標案に明文化しておく必要がある。(審査 WG 委員)
- ・ 性能が向上する場合は受け入れられることをどこかで読んだ記憶がある。性能が低下しない確認を行う必要はあるが、良い方向への変化に関しては、元々想定した変化の範囲で承認してもらえると有り難いと思う。(開発 WG 委員) →評価方法が確立していれば、企業側で個別項目毎に方法を説明できればよいと思う。(審査 WG 委員) →この課題については、企業が一定の基準となるベンチマーク用データを提示し、規定した範囲内で性能が向上する定義を設ける以外に手段がないと思う。新規製品の実用化が則されるように上手く規制していただけたらと思う。(審査 WG 委員) →先生方の意見を踏まえて、薬事的な扱いに関しては、厚生労働省と相談する。(審査 WG 事務局)
- ・ 性能が一律に変わる製品の場合は、アップデート版やバージョンアップ版を1つ作ればすべての製品に適用できるが、インストールした医療施設ごとに性能が変わる場合は、性能や品質担保に個別対応が必要になるとと思われる。その場合、企業として個別対応できる体制を整えることはできるのか？(審査 WG 事務局) →分からないとしか言いようがない。例えば2,000 売れたとなれば、クラウドシステムを選択するしかないと思うが、10~20 製品であれば、個別訪問して対応する体制を取ると思う。そこは製品の品質と販売台数とに応じた対応を行っていくと考えている。(開発 WG 委員) →現時点でも様々なバージョンを使用している顧客がいる。例えば、ある顧客は保守契約を締結しておりバージョンアップするが、別の顧客は保守なしで旧バージョンを継続して使用している。このような状況は実在するので、AI についても同様と思われる。結局は個別に対応する必要があると思っている。(開発 WG 委員) →古いバージョンを使っている、リアルタイムにアップデートされる場合は性能が落ちる可能性もある。企業として、保守契約を締結していないからフォローしない選択肢はないと思うが、その辺りはどう対応するのか？(審査 WG 事務局) →リアルタイムの話はなかなか想像し難いが、クラウド的な環境にしないと、現実的には難しいと思う。(開発 WG 委員) →最初からクラウドを選択すれば薬事的な規制も簡単という側面があると考えられる。この件に関しては厚労省と相談したいと思う。(審査 WG 事務局)
- ・ 評価が困難な点については、技術的問題、コスト的問題、薬事的問題を整理して、本評価指標案に記載してあるとよいと思う。(開発 WG 事務局)
- ・ クラウドとインターネットを対等に記載されていることが気になった。例えばイン

ターネットを通じて何かをダウンロードする形で記載すべきかと思う。また、開発側としては、リアルタイムでの学習が実現可能であることを記載すると共に、その規制についても上手くまとめてもらいたいと思う。(開発 WG 委員)

終了予定時刻を迎えたため討議を打ち切り、開発 WG の進捗状況と今後の方向性確認、並びに審査 WG との作業調整を行った。

<開発 WG の進捗状況>

- ・ 開発 WG としては、評価指標案たたき台の内容を確認し、開発側の対応と課題ついて議論していく予定であることから、現時点では、情報共有するほどの内容はない状況である。(開発 WG 事務局)

<今後の方向性確認及び審査 WG との作業調整>

- ・ 両事務局の相談の結果、12 月 11 日月曜日の会議を合同開催することを考えているが、希望、異論等あれば、各事務局にご意見をいただきたい。(審査 WG 事務局)
- ・ 評価指標案たたき台である資料 6 の議論が完了していないことから、開発 WG 側でも議論して、問題点や懸念事項などを審査 WG へフィードバックする。それを受けて審査 WG でさらに議論し、可能であれば 12 月 11 日の合同会議で方向性を提示できればよいと思う。(開発 WG 委員) → 審査 WG 側では、資料 6 の内容を 11 月頃までにブラッシュアップしてまとめる予定である。それに伴い開発 WG でも 10 月末を目処に意見をまとめてもらい、審査側で適宜取り入れて改訂版を作成し、合同会議で討議したいと考えている。(審査 WG 事務局)

5-3. 審査 WG 第 3 回会議総合討議

資料 6「人工知能を利用した画像診断機器に関する評価指標案たたき台(事務局案)」の修正に当たり、各委員の役割分担を以下のように決定した。

- ・ 1 項：橋爪座長
- ・ 2 項及び 3 項：清水委員、藤田委員
人工知能の定義については、一般的というより、本評価指標案で取り扱う人工知能の定義を記載する。
- ・ 4 項：定型的な文章もあるため、委員からの意見に基づいて事務局で修文する。
- ・ 5 項：全委員から改めて意見を収集し、事務局で取りまとめる。

<質疑>

- ・ 5/(2)/③項に評価すべき項目が列挙されているが、今後新しい手法が開発されると思われることから、記載形式を変更した方がよいと思う。他の項目も箇条書きで詳しく記載するより、もう少し幅をもたせた形式に記載できればよいと思う。(委員)
- ・ 申請時と一変の部分が核になると思われる。(委員)
- ・ 開発側の意見のように、薬事的、技術的、コスト的問題を開発 WG で整理し、審査 WG へフィードバックしたい。(委員)
- ・ 今回のたたき台には、人工知能に特化されていない項目も入っており、且つ従来と違う内容が記載されている箇所があるように思われる。性能が変化しないシステムにも当てはまる場合は、CAD 通知との整合性を考えなくてよいのか。(委員) → CAD 通知に上乘せする観点からすれば、たたき台には不要な部分も含まれている。他にも不必要と考えられる記載があれば指摘して欲しい。(事務局) → 「④ 使用環境」は CAD 通知と整合しているのか？ クラウドシステムという記載はあった

のか？（委員）→CAD 通知にはクラウドシステムという用語はない。この通知は平成 23 年度に発出されており、その時点では単体ソフトウェア自体が医療機器ではなかったことから、装置組み込み型のソフトウェアが対象だったが、一応、単体ソフトウェアもカバーできる仕様としている。基本的には CAD 通知を踏襲しながら、AI に特化した項目を記載することを想定している。（事務局）

- 従来の性能が変化しない製品に対して、CAD 通知では記載がなかったが、必要と思われる事項があれば、本評価指標案に記載する必要があるか？（委員）→今回はあくまでも CAD 通知に上乘せする評価指標案を作成することになることから、記載する不要と考えている。従来の性能が変化しない製品に対して、追加事項が必要であれば、CAD 通知を改正する可能性もある。（厚労省）→ソフトウェア関係の通知を上手く引用する形をとれば良いと思われる。（PMDA）
- リアルタイムでの性能変化を意図する機器に対してはいろいろな意見があるため、深く議論することが必要である。FDA の対応に関する情報を教えて欲しい（委員）→AI に限定した審査は行っていない。ただし、FDA もソフトウェアが頻繁にアップデート又はアップグレードされることは認識している。リスクが低いプログラムについては、例えば製品自体ではなく、メーカーを認定し、当該メーカー限定で申請資料の簡略化や届け出のみで対応可能な仕組みを考えている。（厚労省）→具体的にはどのようなタスクの場合、そのような申請が認められているか？（委員）→最近、パイロットプログラムが始まった。パイロットプログラムに採用された企業には、アップルやサムスンなどの企業のみでなくベンチャーもある。検査数値の分析ツールや糖尿病患者の管理用アプリ等、様々な製品が存在するため、特に画像診断に関連する製品の有無は不明であるが、モバイルアプリのような製品が多い感触がある。（厚労省）
- ベンチャーであればどのような製品開発を行っているか特定できるのでは？（委員）→特定できる（厚労省）
- リスクの高低を判断する考え方に関する情報があれば教えて欲しい。（委員）→FDA の場合は、IMDRF 文書で決めた criticalness と significance の 2 軸で考えている。CAD プログラムで cervical cancer が対象であればクラス 3 に該当する製品もある。（厚労省）

5-4. 次回会議までの流れ

- 第 4 回審査 WG 会議を 10 月 27 日に開催する。そのため、本日より 10 日後を目処に、たたき台に関するコメントを寄せてもらい、事務局で内容を修正する。第 4 回審査 WG 会議では、修正案に基づいて、特にリアルタイムの性能変化に係る事項に関して議論を深める。
- 12 月 11 日開催予定の第 5 回審査 WG 会議を第 2 回合同会議とする件については、開発側と相談した上で決定する。
- 第 6 回審査 WG 会議を翌年 1 月 19 日に開催する。

5-5. 審査 WG 第 4 回会議について

日時：10 月 27 日（金） 14:00 - 17:00

場所：TKP 品川カンファレンスセンター ANNEX カンファレンスルーム 4

以上

次世代医療機器・再生医療等製品評価指標作成事業
人工知能分野 審査 WG
平成 29 年度 第 4 回会議 議事概要

1. 開催日時：平成 29 年 10 月 27 日（金）14:00 -17:03
2. 開催場所：TKP 品川カンファレンスセンターANNEX カンファレンスルーム 4
（東京都港区高輪 3-13-1 TAKANAWA COURT 3F）
3. 出席者（委員：座長以下五十音順・敬称略）
<審査 WG>
座長：橋爪 誠（九州大学）
委員：佐々木 毅（東京大学）、清水 昭伸（東京農工大学）、田中 聖人（京都第二赤十字病院）、中田 典生（東京慈恵会医科大学）、藤田 広志（岐阜大学）、待鳥 詔洋（国立国際医療研究センター国府台病院）、横井 英人（香川大学）
医薬品医療機器総合機構（PMDA）：岡崎 譲、菅原 明彦、加藤 健太郎、谷城 博幸、
今川 邦樹、長谷川 浩一
オブザーバー：坂無 英徳（産総研）、浅沼 直樹（AMED）、岩田 倫明（AMED）、
関口 充（AMED）
事務局（国立医薬品食品衛生研究所）：畠島 由二、中岡 竜介、加藤 玲子
4. 配布資料
資料 1：第 4 回審査 WG 会議議事次第
資料 2：委員名簿
資料 3：第 1 回合同委員会及び第 3 回審査 WG 会議議事概要（案）
資料 4：人工知能技術を利用した画像診断支援システム等に関する評価指標案（事務局案）
資料 5：開発 WG からのコメント集
資料 6：PMDA 提供資料
5. 議事内容
5-1. 開催挨拶及び資料確認
座長挨拶後、事務局から配布資料の確認を行ったと共に、新規参加者が紹介された。
5-2. 前回会議における懸念事項に対する回答
資料 3 を用いて、事務局から以下の説明を行った。
 - ・ 個人情報保護法下で第三者に提供できる医療情報はオプトインに限定されるが、次世代医療基盤法により、使用目的が医療機器の研究開発である場合は、患者の医療情報はオプトアウトで第三者（認定事業者）へ提供可能である。
 - ・ 患者が第三者への提供を拒否した場合、その時点から、当該患者の医療情報は第三者へ提供できなくなる。但し、拒否以前に提供された情報の利活用には影響を及ぼさない。
 - ・ 次世代医療基盤法においても、匿名加工情報について照合して再特定する行為は禁止されている（第 18 条）。よって、認定事業者が管理表を作成することはないと考えられる。

- ・ 希少疾患の場合、データから患者が特定される懸念があるが、次世代医療基盤法においては医療情報が一般に提供されることを想定していない。匿名加工した情報の提供先は、認定事業者と契約を締結した大学や研究機関等に限定されるため、外部に情報が漏れることはないと考えられる。ただし、基盤法の運用時には、当該事項について検討する必要があると考えられている。
- ・ 患者情報の所有権は誰にあるのか？（座長）→ 次世代医療基盤法には、所有権の概念はなく、あくまでも情報の取り扱いを規定している。

5-3. 評価指標案の構成変更に関する報告

資料 4 は、前回会議に提示した評価指標案と比較して項目構成が大きく変更されたため、改定経緯及び内容について、事務局から以下の説明を行った。

<事務局からの説明>

人工知能を利用した医療機器における問題点と解決策を先に示した方が利用者にとって理解しやすいと思われるとの意見があったことから、1 項「はじめに」を(1)項「概説」と(2)項「人工知能を利用した医療機器における問題点と解決策について」に分割し、後者に必要事項を取りまとめ、後述する具体策を参照する形式とした。また、3 項「用語の定義」以降も適宜変更した。主な変更点、並びに関連する質疑応答の内容は以下のとおりである。

- ・ 1/(2)/1)項「ブラックボックス性」：<問題点>診断アルゴリズムを提示することは不可能である。→ <解決策>開発設計段階の学習アルゴリズムやネットワーク構造を示すと共に、確認可能な機器の性能範囲等を規定・確保することにより代替することとした。
- ・ 1/(2)/2)項「性能変化」：予想外の方向に性能が変化する可能性があるため、利点とリスクを勘案した上で、適切な対応策が求められる。
- ・ 1/(2)/2)/①項「変化した性能の検証方法」：<問題点>変化した性能をどのような方法で検証するか？→<解決策>予め市販後に生じ得る性能変化範囲や臨床上許容される性能水準を規定すること等が求められると共に、該当機器に適用可能な検証方法及びそのスケジュール等を定めておくことが求められる。また、使用したデータベースの妥当性を示すことも要求されることも明記した。具体策は後述する該当項目を参照するように記載した。
- ・ 1/(2)/2)/②項「性能変化に伴う品質保証」：<問題点>学習形式に関する要求事項は？→ <解決策>学習方式の明確化、データの詳細とその妥当性を示すことを要求することとした。
- ・ 1/(2)/2)/③項「薬事上の手続きにおける考え方」：<問題点>どのようなケースが届け出、一変、新規申請に該当するのか？→ <解決策>ケース毎の対応例を示しておく必要があると思われる。例えば、バージョンアップに費用を要する場合、ユーザによってはバージョンアップを望まないケースも想定される。その際、旧新の両バージョンが同時に存在することになるが、現行の薬機法で対応できるのか？（事務局）→現存バージョンに性能を付加して新バージョンにする場合、一変申請時に旧バージョンも市場に存在することを明記しておけば、両バージョンの共存が可能である（PMDA）。→ リアルタイムで性能が変化する場合、エンドユーザ毎に性能が異なる機器となる。この場合、性能を確保するための手法について議論しておく必要がある。（事務局）
- ・ 1/(2)/3)項「責任の所在」：<問題点>その使用に伴う責任の所在が曖昧になる。→ <解決策>メーカーと医師等が負う責任について明記するように求めている。具体的な対策は 5 (4) 項を参照するように記載した。

- ・ 3 項「用語の定義」：清水委員担当部分の文献情報（PMDA 科学委員会報告書）に関しては、報告書公開後、記載することとする。
- ・ 5/(2)/2)項「学習」：複数の学習方式を列挙しているが、各々の技術に応じた留意点があれば、その概要を記述した方が良いのではないか？ モダリティから画像撮影装置に表現を変更したが、他によい表現があれば提案して欲しい。（事務局）
- ・ 5/(3)/2)項「市販後」：リアルタイムに性能が変化する製品の場合、性能変化が生じた際にユーザへ通知するシステムを有することが望ましいと記載しているが、開発側からは不要ではないかとの意見があった。性能変化の内容に応じた薬事的な手続きに関する記載は、PMDA からの意見に基づいて修正するが、各委員の意見も収集したい。（事務局）
- ・ 5/(4)項「リスクマネジメント」：不具合対策に代わり新設した。前回までのバージョンには、不具合が発生した際の通知機能を備えた方が望ましいとの記載があったが、WG 委員から、通知先に関する照会のほか、医師の責任により規制当局へ報告すれば、機器自体に通知機能を備えることは不要と思われるとの意見があった。また、対象機器は診断支援が目的であることから、基本的に確定診断の最終責任者は医師と規定することが求められるが、人工知能の判定をもって直接的に確定診断とすることのないよう、適切なユーザトレーニングの実施を要求事項として追加する。（事務局）

5-4. 総合討論

資料 4 の内容を精査した。主な討議事項は以下のとおりである。

- ・ 基本的な性能を担保するため、始業点検としてキャリブレーションを含む作業を行うことにより、最低限の性能を担保できると思われる。また、性能が低下した場合、前の性能レベルにリセットする機能を付加すること等を要求すれば、一定の機能を保つことが可能になると思われる。（委員）
- ・ 性能が常に変化する機器のイメージとして、仮名漢字変換が挙げられる。性能が極端に変化しないことを想定していると思われるが、そのような認識で良いのか工学的観点からの意見を聞きたい。（委員）→ 使用者が現場で学習させることにより、システム精度がリアルタイムに変化するバージョンはいずれ開発されると思われるが、今回の評価指標案を見る限り、対象とするシステムの範囲が不明瞭である。どこまでのシステムを念頭において議論すれば良いのか？（委員）
→ 事務局回答
 - 現時点で考え得る全てのシステムを網羅したい。
 - 従来の「コンピュータ診断支援装置に関する評価指標(以下 CAD 通知)」でカバーできる範囲は、この通知で対応する。ただし、CAD 通知が発出された当時はクラウドシステムが存在していなかったと共に、単体ソフトウェアも医療機器ではなかったため、現状に対応できていない部分もある。
 - 人工知能の能力の程度が未知である。第 1 回会議における講演でも、現時点では、学習のためにデータのチューニング作業が必要であり、リアルタイム学習は実質上不可能であることが紹介された。
 - しかし、他分野ではリアルタイム学習が可能になってきているため、近い将来、医療分野にも当該技術が導入されることを想定し、開発を妨げないように注意を払いながら、リアルタイム学習する機器も包括できる内容にしたいと考えている。
- ・ 企業ヒヤリング時、エンドユーザレベルで常に学習しながら性能が変化するシステムについては、メーカでは責任が取れないとの回答があった。まずは、メーカ側で管理し、定期的にバージョンアップする形式を想定して議論することが重要と思わ

れる。その上で、メーカー側が責任を取れないと考えているシステムについて、どこまで踏み込むかが焦点となる。(委員)

- ・ 前回ヒヤリングは国内企業のみだったが、米国ではリアルタイムに性能が変化することを想定して議論が行われている。(委員)
- ・ システムに与えるデータで性能が決まってしまうため、管理されていないデータを与えた時の結果を判断できない。バッチ式の場合、学習データの選択、使用及びそれらのバリデーションをメーカに求めるよう記載しているが、エンドユーザのレベルで性能変化させる場合にもそれと同等の要求をしないと、医療機器として性能を保証できないと思われる。しかしながら、事後学習の結果、性能が極端に変化しない安全機構等を備えることができれば対応できる可能性があると思われる。(委員)
→ 完全に制御できない製品を評価指標案の対象から外すより、その可能性は残しておいた方がよいと考える。そのための安全機構については、議論の余地がある。(委員)
- ・ 評価指標案には段階に応じた要求事項を記載することになると思われる。(委員)
→ 医療機器として承認するにあたり、一定の範囲内で使用し、一定の条件下の学習データを使用して性能を変化させる場合の要求事項を記述することが求められると思われる。(委員)
- ・ 最終責任を医師に課すのであれば、メーカにあまり負荷を与えない記述にした方が、技術は進化すると思われる。現状では、一定のデータセットを用いて評価する機能を有すること以上の要求は難しいと思われる。(委員) → 今回の評価指標案には、全ての責任は医師にあると記載されているが、機器に不具合が生じていることに気づかずの使用し、間違った判断をした場合も医師の責任になるのか？ PMDA 科学委員会でも話題になったが、ログを残し、検証できるシステムを導入し、検証の結果として機器の不具合が認められた場合は、メーカの責任も問われる可能性を示しておく必要があると思われる。(委員) → 安全性、有効性の確保とトラブル対応において、メーカ側が責任を負うことに繋がるので必要だと思われる。ただ、リアルタイム性と様々な変化に対する問題への解決策を現時点で具体的に記載できるか疑問である。記載するにしても、動的手段で性能変化する機器については、曖昧な表現にならざるを得ない。また、バリデーションの確保に関しては、振れ幅が小さいと考えられる一方、危険な状態となる可能性も想定されることから、現段階で明文化して規制することは困難であると思われる。現時点では、安全性、有効性の確保とトラブル対応に関しては、メーカ側が責任を負うことを記載することが限界と思われる。(委員)
- ・ 現段階で具体的にイメージしやすいのは、「自動運転」である。自動運転中に事故を起こさないためには、AI が最低限のルールを認識し、稼働することが必要であると議論されている。対象機器の最低限のルールを確保する上で、例えばネットワークに接続し、定期的に初期性能を保持しているか確認するためのキャリブレーション機構等が必要になると考える。(委員) → 想定されるキャリブレーションとは何か？(委員) → 初期データセットを用いて正解を導き出すことができるか確認することがキャリブレーションに相当すると考えている。(委員)
- ・ CADe であれば病変の有無の判断になることから、正誤割合のほか、以前のメーカで非検出であったデータが、新規メーカでは正しく判定されているか、或はその逆のケースも判断できる。CADx は病変の有無の判定ではないため評価方法の選定は難しいが、最初に正解であった結果が、性能変化後も正解となることが最低条件になる。例えば、対外診断薬の場合、正解であるべきところを間違うと、不正解を導き出した理由が厳しく検証される。(委員) → CADe、CADx を問わず、キャリブレーション用の基準データがあれば問題ないと思われるが、ROC 曲線が交差する

場合は、性能の向上と低下を判断することが難しくなる。評価の良否基準は感度等の単一パラメータでないことが多いことから、どのような観点で判断するか難しいと思われる。(委員) → ROC 曲線で面積が一緒でも、用途に応じて、特異度と感度のどちらを上げるかを承認申請書に記載しておけば、判断の方向は定まると思われる。(委員)

- ・ メーカーがバージョンアップする場合に使用する学習データの保証や性能の担保について規制があれば、オンサイトで学習させる場合も同じ規制の下に対応しないと矛盾が生じる。ただし、エンジニアでない医師等がその要求に応えるのは難しいと思われる。(委員) → 機器に自己診断機能を持たせ、性能が低下する場合、変更部以外を維持することで結果的に機能低下を防ぐ仕組みをメーカーが作ることができれば対応可能ではないか？ オンサイトでも学習データに関するログは残す必要があり、工夫次第で学習データの保証は可能ではないか？(委員)
- ・ オンサイトの場合、メーカーが準備したデータセットに新規データが蓄積されていくことになる。それをどのように臨床で使用するか整理しておきたい。現状では、データを上乗せ又は再学習させる場合ともに、研究の範疇になると思っている。そのような機器を臨床で使用する可否かは医師の判断になり、適用外使用や未承認機器使用の場合と類似した状況になることから、医師が相当の覚悟を持って使用しなければならないと思われる。そのため臨床現場には、性能が担保された医療機器と研究レベルの機器が併存することになると思われるが、その点についてどう考えるか？(委員) → 常に新しいデータを個人レベルで追加することはリスクが高いため、現状では、企業側が準備した学習データや教師付きデータを使ってバージョンアップすることを想定していると理解している。(委員) → その次の段階として、オンサイトで学習させて性能が変化する機器に対する対策も考慮しておくべきだと思われる。(委員)
- ・ 1/(2)/2)/②項「性能変化に伴う品質確保」に記載されている「機器から取得したデータを学習に用いる場合の妥当性判断には、現時点の科学水準に鑑みて専門知識を有する医師による確認が必要と考えられる。」に関しては、完全に制御できる環境で新しく学習するシステムとオンサイトで学習するシステムを分けて記載しておくべきだと思われる。(委員)
- ・ 米国におけるリアルタイム学習に関する対応の現状を確認する。(委員) → リアルタイム学習の研究の方向は否定しないが、評価指標案に記載する対象範囲が問題である。例えば、アメリカでデータの妥当性判断を医師が行うことを要求されるか現時点ではわからない。両国におけるルールの違いを理由として受入れないことが起きないように、柔軟な表現にしておいた方が良いのではないか？(委員)
- ・ 1/(2)/2)/②項「性能変化に伴う品質確保」に関して、具体的にどのような表現で記載すればよいか提案して欲しい。一般の医師が病院で使用し、学習データとしての妥当性を判断することは難しいと思われる。例えば、そこにメーカーが関与せざるを得ない状況が起こり得るとも考えられる。(事務局)
- ・ PMDA から、資料 6 を用いて市販後の性能変化に係る対応案の説明があった。
 - 資料 6 は現状を整理するための資料であり結論ではない。
 - 想定される状況を鑑みて、2 パターン考えた。
＜パターン 1＞：アップデートをメーカー又は病院が行ったでも大差がないと考えられる場合
 - 市販後アップデートをどのタイミングで誰が行うかを場合分けし、どのような違いが発生するかとの観点で作成してある。
 - A：標榜や仕様を変更する場合は、メーカーがバリデーション試験を行い、一変申請することになり、従来の医療機器と相違がないため、今回議論する必要は

ないと思われる。

- **B**：データが一定数収集される又は入手される度に、性能をアップデートする場合については議論の余地がある。メーカー側又は病院側でのアップデートを問わず、市販後も臨床上仕様できる水準を維持し、その範囲内で変化する場合の問題ないと思われる。そのためには、市販前にアップデートするための条件や頻度等を全て規定し、性能評価方法についても説明を求める必要があると思われる。例えば、アップデートの実行可否を判断するテストを実施する。その際、テストデータの妥当性のほか、誰がどのように用意するか考える必要がある。

＜パターン 2＞：使用者毎の学習（オンサイト学習）とサーバー上での一括学習（オフサイト学習）の場合

- オンサイト学習

特 徴：使用者毎にアップデートの可否を判断できる（アップデートしない選択も取れる）。また、使用者毎に性能の調整ができる（モダリティの差異等の吸収）。

問題点：使用者毎に適切な教師付き学習データが入手できるか懸念される。病院固有の診断の癖を学習すると、その病院で見逃し易い症例は結局見逃す可能性が高い。

対 応：使用者が適切に学習させることができる情報提供や体制構築が必要である。アップデート時の試験結果を解釈できる人物の配置が必要となることも想定される。

- オフサイト学習

特 徴：性能の変化を企業が把握・管理できる。

問題点：使用者が意図しないタイミングで性能が変化する。

対 応：性能変化が発生したこと、並びに性能変化の影響について、使用者へ適切に情報提供することが必要である。

＜補足コメント（PMDA）＞

規定された学習プロセスから逸脱して得られたデータを学習に使用することも想定されるが、その場合には、必ずその検証が必要となる。

オンサイト学習で性能を変える場合、ソフトウェアや AI だからこそ、B のような考え方で、上述したデータの検証方法も含め、性能を確保するためにメーカーが取るべき手法が明確に規定されていれば、病院側で学習可能とする形でまとめられるのではないかと考えている。

- ・ 使用者でアップデートする場合の条件を規定することは可能か？（委員）→全ての学習データを管理・評価して性能が下がっていないことを確認することとすれば良いと思われるが、一つの実現可能性として、メーカーが検証手法を創意工夫して作成し、その妥当性を判断することが考えられる。（委員）→ 薬事手続きは未定であるが、B の場合でも、メーカーが情報を収集し、定期的に報告するシステムを第三者が確認することが必要と考えられる。既存医療機器と比較した場合、より先端的な体制が求められることになるが、そのようなフレックスな考え方は可能であると思われる。（PMDA）
- ・ ソフトウェアや AI でも、異なった目的に使用した時はどうなるのか？（委員）→従来の医療機器と同様、予め定めていた適用範囲を超えていると判断される。しかしながら、オンサイトで使用者が性能を確保する場合、医師の判断の範疇で責任を取ると言われれば、性能の向上又は低下を問わず、メーカーに責任を課すことなく、臨床現場で使用することが認められる可能性はある。ただし、製品の性能を上げるために使用するデータが学習に値するデータであるか判断する手法やその評価法等を規定し、倫理承認を取得することが求められると考えられる。なお、病院側で

データの質を評価・判断するために必要な現時点で想定される手法等を承認申請書に具体的に記載すれば、オンサイトでの学習を含んだ承認が可能になると思われる。(PMDA)

- ・ オンサイトとオフサイトで分けて記載する。オンサイトの場合、コストを考慮しなければ、メーカーと同じレベルで性能変化を管理することができると思われる。具体的な内容はメーカー側から提案する余地を残した記載にすれば、海外製品にも対応できると思われる。(委員)
- ・ オンサイトで性能変化が生じた場合に必要な薬事手続きは何か？(委員) → 承認された使用方法・目的の範疇に該当するか否かで判断されることになる。(PMDA) → 性能を担保するための方法をメーカーがどのように記載するか、本評価指標案の中で示すことが鍵になる。(委員) → メーカー側は、一括学習、バッチ学習時と同等レベルの保証方法を記載することになるのではないかと？(委員) → そこまで要求すると、オンサイトのハードルが上がり過ぎてしまう。(委員) → 同等をどのレベルまでと考えるかによるが、そこまで求める必要があると思われる。(委員) → オンサイトとオフサイトでは更新のタイミングが全く異なることから、同等と記載すると厳しすぎると思われる。例えば、定期的に販売時のデータセットを用いて性能を確認することを義務付けること以外、現時点では記載できないと思われる。(委員)
- ・ 学習させることを前提とした適用が問題になると思われる。性能に関しては、学習後必ず上がる保証がないため、性能変化範囲を規定した標準機能として薬事承認を取得する。その後は、オンサイト先の責任で運用することになると思われる。オンサイトにおける性能の振れ幅はある範囲に収まると思われるが、性能の向上・低下を問わず、サイトの責任で性能変化させても規定された性能範囲を越えなければ良い。最終的な性能が規定範囲を超えてしまい、リセットが必要となる場合は標準機能に戻すことができるようにする等、サイトで学習させた結果の性能と標準性能を区別できるようにしておくが必要になる。そのような機器を薬事承認するためには、厳密性と柔軟性が必要になると思われる。(委員)
- ・ オンサイトで性能が徐々に改良される使い方をする場合、届出や一変申請等が必要なのか？(委員) → 今回の場合、従来の医療機器と異なるため画期的な承認方法になるが、性能変化を含めた承認とすることをイメージしている。逆に言えば、承認された範囲内での変化であれば、届出も必要ないと思われる。ただし、サイト毎の性能変更にはある程度の制限を付すことになる。(PMDA) → 使用法を一括して申請すれば、あえて数値目標まで記載する必要はないということか？(委員) → 今後議論が必要と思われるが、例えば、ある感度・特異度を超えた時のみアップデート可能とする場合、それは具体的な数値条件になると思われる。(PMDA)
- ・ 資料6のパターン1では、オンサイトで自由に学習しているが、定期的にメーカーの関与があり、販売後もメーカーの管理下にあるとイメージしている。(PMDA)
- ・ 従来の医療機器は静的な承認手法であるが、AI に関しては動的な承認手法を認めるための必要十分条件が記載されていれば良いというイメージを持っている。つまり、動的に性能が変わり得る範疇を含めてどのような条件を課せば、動的に管理できるかが鍵になると思われる。(PMDA) → その範囲を超えた時のみ、一変申請が必要となるのか？(委員) → PMDA としても経験がないので、資料6に示したように様々なパターンを想定しているが、現在も PMDA 内で議論しているところである。(PMDA)
- ・ オンサイトでの性能変化は病院毎に異なり、全てを管理することは難しいため、メーカーからもそのような機器の開発には消極的な意見があったと思うが、実際には大幅な性能変化は求められていないように感じている。病院毎に判断して学習させる

場合、各病院にスペシャリストができるというイメージがある。一方、一括で学習させる場合は、メーカーからスペシャリストが派遣されるイメージとなり、診断能力をメーカーが保証していると思っている。この点、委員の先生方はどのように考えているか？(PMDA)→ 疾患群や対象とする病変は病院毎に全く異なっているため、その疾患等に特化した機器へ変化する可能性がある。それらの中から、性能が突出したオンサイトデータを収集して事後学習させることで、総合的なレベルも上がる利点がある。(委員)

- ・ 内視鏡がネットワークに接続されて機能するのは将来のことになると思われることから、柔軟性をもった記述にしておくのが良い。(委員)→ 方針としては同意するが、性能テストに関しては厳密にしておく必要がある。性能の振れ幅は ROC で計算することができるが、メーカーによっては、ROC を使用せずに、広範な振れ幅を設定して、市販後に手続きをしなくて済むように申請してくる可能性がある。(委員)→ 振れ幅に関しては、CAD 通知に従えば良いと思われる。但し、その振れ幅を必要以上に大きく取る懸念はある。(事務局)→ その振れ幅も、CAD 通知である程度定義されており、性能評価を統計学的根拠に基づいて実施するよう記載されていたと思う。(委員)→ 審査の立場からすれば、疾患や位置付けで数値は変わると考えている。仮に極端に大きな振れ幅で申請されたとしても、その根拠を示すよう求めることになる。AI の技術面から基準となる数値等があれば、本評価指標案で示して欲しいが、臨床上の評価であれば、CAD 通知に従えば良いと思われる。(PMDA)
- ・ 転移学習ではソースが不明瞭であり、コントロールし難い状態になっていく可能性が高い。使用したデータの信頼性を別に判断する必要があると思われる。(委員)→ ログを残す必要がある。(委員)→ 個人情報保護法の中で、データを再結合しないということであったが、変換テーブルを持つことがリスクになる反面、変換テーブルがないとデータソースの検証ができなくなる問題がある。(委員)→ ソースの検証という意味では、連結で遡及できなくても構わない。現時点でも放射線や病理分野では、画像のタグ付けとしてテキストデータを持つデータベースを考えている。この形式を採用すれば、提供側の病院に遡及しなくても大丈夫である。ただし、画像の蓋然性や情報のログを取って検証できる仕組みを持つておくことは必要である。(委員)
- ・ 医学系研究に関する倫理指針等を含め、様々な法的規制が制定されている中、匿名化した個人情報の取り扱い方を担保するのが認定業者の役割であると思っているが、今後、大元のデータのラベリングポリシーが違っていたため、後から加えることができないデータであったケースが想定し得ると思われる。全ての過程において、同じレベルで、同じポリシーでラベリングされている訳ではないことを念頭においておく必要がある。(委員)→ 申請時に情報ソースを明記できれば良いが、将来的には、タグのみが付いた状態でランダムにデータを集めてくる可能性が高い。そのため、様々な状況を想定して、データソースをどう担保しているか判断できる要件が必要になると思われる。(委員)
- ・ 病理医が少ないことから、AI を利用した病理自動診断支援ツールの開発が考えられている。この支援ツールを使用するのは、病理医がいない病院であることが多いと考えられるため、使用者は病理医以外の医師となり、得られた結果が実際の組織を確認することなく採用される可能性がある。そのため、病理分野でのオンサイト学習において、間違っていると思われるデータがある程度揃わないと、診断支援結果が間違っていることに気付かないことが生じ得る。結果として、そのデータが集積するまでの間、被害を受ける患者が出てしまうことが懸念されるため、病理自動診断支援ツールに関しては、学習に関して、病理医の指導下に学習を行う条件を付

す必要があると思われる。(委員) → この問題は AI のみでなく、他の医療機器にも当てはまる問題である。例えば、外科医が使用する手術ロボットを内科医が使用しても上手く使えない。(委員) → 内科医が外科医用の機器を使用することは、ほぼ無いと思われるが、日本の場合、病理医が 2,400 人程度しかいないため、AI を利用した病理自動診断ツールが開発された場合、使用したい病院は多くあると思われる。(委員) → 本評価指標案が発出された後、その機器は自動診断の範疇になるか？(委員) → 病理学会で開発を目指している「病理自動診断支援ツール」は、病理医がいない病院でも自動で使うことを目指しているが、それが実現した場合のリスクを懸念している。(委員) → 自動診断は消化器内視鏡でも考えられている。支援装置になると思われるが、実際はほぼ自動診断装置である。(委員) → その場合、オンサイト学習によりアルゴリズムに変化が生じた場合、ブレーキが効かないと思われる。何人かの患者に実害が出て初めて、ブレーキがかかることになる。病理組織の場合、アルゴリズムやパラメータの少しの変化で、正解率が大きく変わることがあるため、慎重に行う必要がある。病理診断は最終診断に直結するため、基準を詳細に規定する等のハードルを考慮して欲しい。(委員) → 第 1 回会議において、個別分野の評価項目については深く踏み込まないこととしたが、必要であれば付録に記載することは可能か？(委員) → 過去の評価指標では、特定機器の概要と留意点を記載した別添を付けた経緯はある。もしくは、最後に新たな項目を追加し、個別分野の製品に係る留意点を特出しして記載することも可能と思われる。(事務局)

- 映像や画像のみでなく、診断名等のタグを付け、検証することが重要と思われる。ログの残し方、学習用教師データの選別機能に関する基準を示すことができれば、この問題はクリアできると思われる。(委員) → オンサイトの場合、教師付きでリアルタイムに学習することは、ほぼ不可能と思われる。(委員) → オンサイトでは教師付きデータを使用して学習すると考えていたため、性能が病院毎に特質化することを想定していた。教師付きであれば、特化したデータセットが採用出来る等、幾つかイメージできるが、教師なしの場合の終着点については議論が必要である。(委員)
- 病理は最終診断になることから間違えは許容できない。内視鏡の場合、オンサイトで胃カメラ検診を行うことは想定されるが、非常にハイリスクな患者の内視鏡での診断を AI に委ねることはあり得ないと思われる。病理診断の場合、オンサイトで教師なしに性能を変化させる製品は規制せざるを得ないと思われる。現時点の技術で教師なしデータでオンサイト学習を取り上げる意味が理解できない。オンサイトで集めるデータはログ及び教師付きにするべきと考える。(委員) → AI 技術の優れた点を医療機器に利用する場合、少数のオンサイトでの症例のみでは、AI の機能自体が発揮できないのではないかと。全国規模のデータを一括して常に拾い上げるシステムが構築できる状況下で、リアルタイムに性能が向上する機器が開発されれば、AI としての機能が発揮できると思われる。(委員) → オンサイトで大量のデータをサンプリングするのみでは、教師なしデータが増えてくることになる。ログを残すことにより、データが「教師付き」となるので、それは意味がある。そうすれば、複数の病院のデータを合わせた研究も進めることができる。(委員)
- オンサイトの場合の対象と使用法を特定領域に限定しなくても、医師が見落としした責任は機器ではなく医師にあることが明示されれば、医師も気軽に使用しようとは思わない。(委員)
- 山口大の間普先生が、初期は教師付きデータで学習させ、その後、教師なしデータを追加していき、性能を上げていく研究を行っている。今後、どのような性能を備えた機器が開発されるか未知であるが、数年以内に大きく現在の常識を覆す可能性

があることを想定して、現在の科学レベルの範囲のみに留まらず、様々な余地を考慮しておいた方が良いと思われる。(委員) → 胃がん検出と胃潰瘍検出は機能が異なる。技術が高度化されると、CAdE と CAdx は区別がつかなくなってくる。形態学的な異常を検出する場合は、教師なしデータを大量に取り込ませれば良いが、何を学習させるかにより教師データの定義も変わってくる。本 WG で議論すべき事項は、個別技術の議論ではなく、薬事規制の手法であり、それをどのように記載するか重要となる。(委員)

- ・ 現行の評価指標案では、深層学習という言葉に統一性がない。また、ブラックボックス性という言葉は専門用語なのか？(委員) → 専門用語ではない。内部がわからないことを表現したものであることから、適切な言葉に置き換えれば良いと思われる。(委員)
- ・ 市販前後における性能評価方法が記載されていれば、学習データに関する様々な問題を抱えていても、評価指標案として発出できると思われる。先程までの議論を踏まえて、性能評価をどう考えるべきかについて、より深い議論できれば良いと思う。(委員)
- ・ オンサイトとそれ以外を区別して記載しているのか？(委員) → 現在のバージョンでは、リアルタイムで変化するケースはクラウドしか想定しておらず、バッチ式で性能が変わるケースのみ区別したつもりである。(事務局) → クラウドはデータの保持形態に係ることであり、学習をどこで行うかを区別するのみで良いと思われる。(委員)
- ・ クラウドの反対語はオンプレミスであり、サイトにあるコンピュータを指す。インターネット上に存在する場合は全てクラウドに該当する。また、クラウドでも、個々に使用するケースもあることから、10 ユーザが存在すれば、それぞれが個別に教育されるケールもあり得る。クラウド＝メーカー一括管理とはならないので、誤解を招かないように記載を変更する必要がある。(委員)
- ・ 性能評価のポイントの一つは性能の範囲となる。今の記載では、性能範囲を誰がどのような基準で何を考えて決めれば良いか判断できない。範囲ありきで記載されているため、その範囲の決め方を明示すべきと考える。例えば、統計学的に決める、あるいは一般的な医師の能力を基準にする等、説明可能な形式とすることはどうか？(委員) → それは個々に異なるため、数値化は難しい。(座長) → 医学的根拠、あるいは統計学的根拠を明らかにすることを求めれば良い。(委員) → 承認申請時に性能変化範囲を規定することが大事になると理解した。その場合、事前に規定した性能変化範囲の根拠を示すことを求める等、市販後における性能変化の検証方法を具体的に記載する方が良いと考える。また、PMDA が示したように、メーカーの関与についても記述するように修正する。(事務局) → 必ずログを取ることと、検証データのソースも明記するようにした方が良い。(座長)
- ・ リスクマネジメントの項に市販後の問題に対する企業の責任が集約できると思われる。今回の機器に関しては、企業が想定していない情報をユーザが学習させることで問題が生じ得ると考えられるため、本項に企業の責務、例えば事前コンサルテーションやユーザ教育、特にデータに関するリスクマネジメントを追記した方が良いと考える。(委員) → オンサイトに関する事項は、その他と分けて明確に記載すべきだと考える。(委員) → ネットワークに接続されている AI の場合、特定の病院で機器の性能が変化すると、連動して他の関連病院の機器にも性能変化が生じてしまうことになるが、最悪の場合、その性能が一斉に低下するリスクがある。このようなリスクについても、記載できれば良いと思われる。(PMDA)
- ・ 現行案では、3 項で定義している用語が、2 項で既に使用されている。用語の定義は先にあった方が良いと思われる。また、評価指標案には、再学習、追加学習、自

己学習と 3 つの表記があるので統一した方が良い。(委員)

- ・ 再学習によって性能が変化する場合、データが変わることにより変化が生じる。データは撮影装置の条件変更や母集団が変化すること等によって変わる可能性がある。何による変化を主として考慮するかを評価指標案の最初で述べておいた方が良いと思われる。(委員)
- ・ 現行案のまとめ方は読み難い。(委員) → 問題点と解決策を先に記載した方が良いという示唆があった。(座長) → 要点のみを取りまとめて書いた方が良い。(委員) → 修正を検討する。(事務局)
- ・ 分野毎に別添を付す案があったが、得られた結果が診断に直結する場合、特段の配慮が必要となる観点から記載するのが良いのではないか？(委員) → 病理診断の場合の懸念は、その機器の位置付けに応じた懸念と思われる。例えば「機器の使用目的がスクリーニング的な位置付けの場合は、多くの患者を診察しないと問題点が見えてこないことから、特段の配慮が必要」等の記載の方が分かり易いと思われる。(PMDA)
- ・ 専門家が確認できない場合の観点を追記した方が良いと思われる。(PMDA)
- ・ 同じ医療機器の場合でも、病理医が在籍する病院はオンサイトで学習させても良いが、病理医不在の病院では専用プログラムを介して使用する制限を付す 2 段階構えとすることは可能か？(委員) → 品目を分けるかどうかの議論になると思われる。使用環境等を規定することはあると思うが、品目全体に係る事項であるため、リスクレベルに違いがある場合は、その品目を別品目として取り扱うべきか議論する必要がある。(PMDA)
- ・ 病理に関する留意点は非常に重要と思われるので、別項に簡単にまとめて記載した方が良い。(委員) → 佐々木委員に原案作成を依頼し、事務局で適宜編集する。(事務局)

5-5. 次回会議までの流れ

- ・ 第 5 回審査 WG 会議を 12 月 11 日に開催する。そのため、本日より一週間を目処として、評価指標案(資料 4)に関するコメントを寄せてもらい、事務局で内容の修正を行う。この作業を 2 回繰り返して精査した評価指標案を 12 月 11 日以前にメール配布する予定である。第 5 回会議及び第 2 回合同会議では、当該評価指標案更新版をさらにブラッシュアップする。
- ・ 第 6 回審査 WG 会議を翌年 1 月 19 日に開催する。

5-6. 審査 WG 第 5 回会議及び第 2 回合同会議について

日時：12 月 11 日(月) 16:00 - 19:00(時間配分は後日連絡)

場所：AP 東京八重洲通り 11 階 会議室(予定)

以上

A 市販後の性能は変えない

B データが一定数収集される又は入手される度に、性能をアップデートする。
 → 学習には時間がかかるので、現状マシン性能では、前者が精一杯か？
 → 将来的にマシンスペックが改善されたら、データが入る毎に評価するか？

	変更頻度	アップデートする人	性能を確保する方法	第三者による確認
A	標榜や仕様を変えたいとき	メーカー	メーカーによるバリデーション試験	一変等
B	データが一定数収集されたとき、又は入手されたとき	メーカー 使用者(病院)	- アップデートするデータの条件・頻度を規定 ※1による確認	メーカーが収集し市販後定期報告

※1

- アップデートを行うか否かを判断するテストを実施する
 - テストデータの種類
 - テストデータ全体に対する性能評価(F値、損失関数等)
 - 絶対に判定ミスしてはいけないデータ等あれば、これらに対する評価

【懸念】

テストデータを誰が用意するか、どうやって用意するか、どのようなテストデータならば適切か。

	オンサイト学習	一括学習
学習者	使用者(病院)	製販業者
特徴	• 使用者毎にアップデートの可否を判断できる(アップデートをしたくない病院は学習させない)。 • 使用者毎に性能の調整ができる(モダリティの差異等の吸収)。	• 性能の変化を製販業者が把握・管理できる。
懸念	• 病院毎に適切な教師付学習データが入手できるのか？(病院の診断の癖を学習すると、その病院で見逃しやすいものは結局見逃しやすい)	• 使用者が意図しないタイミングで性能が変化していく。
対応	• 学習者が適切な学習ができるような情報提供や体制の構築が必要。(アップデート時の試験結果の解釈ができる人物の配置？)	• 性能変化が発生したこと、及び性能変化の影響について、使用者に適切に情報提供することが必要。

次世代医療機器・再生医療等製品評価指標作成事業
人工知能分野 審査 WG
平成 29 年度 第 5 回会議 議事概要

1. 開催日時：平成 29 年 12 月 11 日（月） 16:00 - 17:08
2. 開催場所：AP 東京八重洲通り 13 階 A 会議室
（東京都中央区京橋 1-10-7 KPP 八重洲ビル）
3. 出席者（委員：座長以下五十音順・敬称略）
座長：橋爪 誠（九州大学）
委員：佐々木 毅（東京大学）、清水 昭伸（東京農工大学）、田中 聖人（京都第二赤十字病院）、中田 典生（東京慈恵会医科大学）、藤田 広志（岐阜大学）、待鳥 詔洋（国立国際医療研究センター国府台病院）、横井 英人（香川大学）
厚生労働省（厚労省）：青柳 ゆみ子
医薬品医療機器総合機構（PMDA）：岡崎 譲、菅原 明彦、加藤 健太郎、桐山 瑤子、古森 亜矢、宮崎 生子、今川 邦樹
事務局（国立医薬品食品衛生研究所）：靱島 由二、中岡 竜介、加藤 玲子
オブザーバー：鎮西 清行（産総研）、坂無 英徳（産総研）、扇谷 悟（AMED）、浅沼 直樹（AMED）、岩田 倫明（AMED）
4. 配布資料
資料 1：第 5 回審査 WG 会議議事次第
資料 2：委員名簿
資料 3：第 4 回審査 WG 会議議事概要（案）
資料 4：人工知能技術を利用した画像診断支援システム等に関する評価指標案（事務局案）
5. 議事内容
 - 5-1. 開催挨拶及び資料確認
座長挨拶後、事務局から配布資料の確認を行った。
 - 5-2. 評価指標案の構成変更に関する報告(資料 4)
第 4 回会議に提示した評価指標案の項目構成が変更されたため、改定経緯及び内容について、事務局から以下の説明があった。

＜事務局からの説明＞
 - 1 項：「人工知能を利用した医療機器における問題点と解決策について」を 5 項に移動させて、イントロダクションのみとした。
 - 2 項：下記の通り、対象となる機器の機能を具体的に記載した。
 - ・ 疾患名まで特定せず、病変の疑いがある部位のみ検出する
 - ・ 病変の疑いがある部位の検出に加え、疾患名の候補を提示する
 - ・ 疾患名候補の提示に加え、それらの重み付け（順位づけ）まで行う
 - 3 項：タイトルを「用語の定義・説明」に変更し、機械学習の部分を機械学習と深層学習に分割した。また、説明部分を「1) ブラックボックスとしての性質」、「2) 継続的な性能の変化」、「3) データベースの信頼性の問題」としてそれぞれに補足説

明を加えた。

5 項：新たに項立てし事務局がさらに加筆した内容確認を委員に依頼した。

- ・ 5/(2)/2)項「性能変化に伴う品質確保」に、前回会議で提案された「キャリブレーション」、「リセット機能」を支援システム等に備えることを反映させ、5 項、6 項及び 7 項に関連する留意点を適宜追記した。
- ・ 5/(2)/3)項「薬事上の手続きにおける考え方」に、承認後の医療機器プログラムの変更に伴うケースをリスクの大きさの順に想定して列挙した。具体的な手続きはケース・バイ・ケースになると思われることから明記していない。「教師なし」データに関しては、記載の必要がないとの意見もあったが、将来そのようなデータを利用した支援システムが開発される可能性もあることから記載を残すこととし、これに関連する留意点を 7 項に追記した。
- ・ 5/(2)/4)項「責任の所在」には、基本的に支援システム等の使用における責任者は医師と明記することを考えている。

6 項：議論及び確認が必要な点について、以下の説明があった。

- ・ 6/(2)/1)項「原理（処理アルゴリズム）」において明示が要求される「開発設計時のプログラム」は概要ではなくより具体的な記載が必要であるとの意見があったが、どこまで詳細な内容の明示を求めるかについては討議が必要である。
- ・ 6/(2)/2)項「学習」に、転移学習も考慮する必要があるとの意見を受けて、注 2 として転移学習に関しての留意点を記載した。
- ・ 6/(2)/3)項「使用環境」は、今回の評価指標がコンピュータ診断支援装置に関する評価指標への上乗せということから必要ないとの意見もあったが、当該評価指標発出後に発出された通知を引用していることを考慮して残した。
- ・ 6/(2)/4)項「併用する画像撮影装置の特定」は、併用装置によって学習に使用する医用画像データの撮像条件や教師データの作成方法等が異なる場合が考えられたので事務局案として記載しているが、より具体的な記載が必要であれば意見を収集したい。
- ・ 6/(3)項「安全性、品質及び性能評価」は、オンサイトでもメーカと同等の性能評価を要求すべきとの意見があったことから、6/(3)/2)項「市販後」の部分に性能変化の実現方式毎に場合分けして留意すべき事項を記載した。
 - ① バージョンアップ毎の段階的な性能変化を意図する場合（メーカがバージョンアップさせる場合）：性能変化内容に応じた手続きを行うことを記載している。
 - ② バージョンアップ毎の段階的な性能変化を意図する場合（使用者がバージョンアップさせる場合）：医療機関に対して事後学習責任者の指名や手順書の整備等を含めたマネジメントシステムに関する要求事項を追記している。
 - ③ 臨床使用に伴った自律的且つ連続的なバージョンアップを意図する場合：自己テスト機能やリセット機能についても記載している。
- ・ 6/(4)項「リスクマネジメント」は、リスクマネジメントを考える上でネットワークに繋がっているか否かが重要であるとの意見があったことから、それを前提とした記載に修正した。また、不具合によるリスク軽減のために下記の機能を要求事項として追記した。
 - 自己テスト等により使用開始前に毎回性能評価を行い、最低限の性能を確保する機能
 - リセット機能
 - 不具合通知機能
 - 不具合発生時に対象支援システム等を停止させる機能

7 項：前回の議論に基づき、「機器の特性に応じた追加留意事項（案）」として新しい項

を追記した。

- ・ 7/(1)項「事後学習による性能変化が診断支援においてハイリスクとなる場合」に、病理分野の例を挙げその留意点を記載した。
- ・ 7/(2)項「支援対象システム等の対象画像が動画の場合」に、観察像が動画となりうる消化器内視鏡機器や超音波画像診断装置に対する留意点を記載した。
- ・ 7/(3)項「教師なし」データによる学習が可能な機器の場合」に、「教師あり」データを用いた場合よりハイリスクになることから、リスク回避のための要求事項を記載した。

これらに加え、性能変化等に関する修正案全体の内容確認を委員に依頼した。(事務局)

5-3. 総合討論

資料4の内容を精査した。主な討議事項は以下の通りである。

- ・ 本評価指標の対象は、市販後に性能が変化しない機器は対象としないということでもいいか？(委員)→その理解で問題ない。(事務局)
- ・ 3/(4)項「深層学習」の定義が不十分なため、その記載については清水委員に願う。(委員)
- ・ 再学習だけでなく、フィルタ追加等、ニューラルネットワークの軽微な変更により性能が変化する支援システム等も本評価指標の対象とし、届け出なしで性能変化を認める方針だったと記憶しているが、それで間違いないか？(委員)→ニューラルネットワークの些細な変更は性能に対する影響は非常に小さい等の表現に留め、薬事手続きの詳細は本評価指標には記載しない方がよいとの意見が出ているところである。ニューラルネットワークの軽微な変更による性能変化については軽微変更届という手続きにした方がいいという意見なのか？(PMDA)→薬事的な問題を提起してはいない。ニューラルネットワークの構造が軽微に変更されることで、学習アルゴリズム及び性能も変わることが明らかになっている。それについても「性能変化」する支援システム等として本評価指標の対象となることを示すべきではないかと考えている。(委員)→軽微な変更であっても、その変更が性能や安全性に及ぼす影響まで軽微になるかは個別判断になる。(PMDA)→そのような性能変化に関しても、本評価指標の対象に含めることは可能であるため、追加文章の検討を委員に依頼する。(事務局)
- ・ 5/(2)/4 項「責任の所在」だが、医師法の範囲では医師が責任を負うという理解で良いと思われる。ただし、保守やトラブル対応に関してはメーカーが責任を負うべきである。(委員)
- ・ 6/(2)/1項「原理(処理アルゴリズム)」における開発設計時のプログラムに関しては詳細な説明が必要と思われる。(委員)→開発設計時のプログラムとは、主に学習アルゴリズムのプログラムと診断用ニューラルネットワークに関するプログラムの2つが考えられる。記載内容に関しては、注2に記載されている「人工知能技術の専門家が理解できるレベルでその詳細を示すこと」であることが望ましいと考えられる。(委員)→ソースコードまで要求するのか？(委員)→全体を俯瞰できるように説明が少なくとも必要である。(委員)→中身が文章、あるいはフローで特定できるレベルで示されていれば構わないと思う。(PMDA)→概要の中で大事なものは学習アルゴリズムと診断用ニューラルネットワークである。(委員)→本評価指標の該当部分を下記のように修正する。(事務局)

・ 設計時プログラムの概要

➤ 開発設計時の学習アルゴリズム

➤ 開発設計時の診断用ニューラルネットワーク構造

- 6/(2)/2)項「学習」におけるデータベースだが、学会がボランティア的にデータベースを構築している場合もあると思われる。本評価指標にデータベースに関しての要求事項が例示されているが、学会ベースの場合、事業計画書や業務内容等が整備されているとは考えにくい。すべてを記載する必要はないと思うが、どう対処すれば良いか？（委員）→あくまでも例示であり、審査側が必要とする項目が適宜必要となる。（事務局）
- 6/(4)項「リスクマネジメント」に関して下記の意見があった。
 - 「内部データを利用したキャリブレーション」に関して、内部データが医療機関の内部データなのか当該機器システムの内部データかわかりにくいので、「当該システムの内部データ」と変更してはどうか？（委員）
 - 「不具合通知機能」とあるが、不具合が明確でないためその定義が必要だと思われる。また、不具合がどの程度発生するかわからないが、そのような通知が規制当局に大量に届いても対応できるのか？（委員）
 - 「不具合発生時に対象支援システム等を停止させる機能」とあるが、不具合の定義が曖昧な状態だと、ささいなことで止まってしまうシステムを企業が開発してしまうことが懸念される。（委員）
 - 「使用者は専門知識を有し且つその機器のリスクに応じた当該支援システム等に関する教育を受けた医師を始めとする医療従事者に限定する」とあるが、現在学会連携でユーザトレーニングを行っているソフトウェアもあることから、AI も同様の形式になると思われる。要求するユーザトレーニングの前に「関係学会等と連携し」を挿入してはどうか。（委員）
- 7/(3)項「医師の代替能力」とあるが、人工知能のよって医師が代替されるようなイメージを持ちかねないので、削除するか別の表現に変更した方が良い。（委員）
- 国内メーカーが国内外の AI 業界で不利にならないように、臨床系の学会で使用する医師の養成コースを設ける等の形で AI システムをオーソライズすることも必要だと思われる。学会レベルで医師の質や AI リテラシーを担保することも必要になると思われる。（開発 WG 事務局）
- 北米放射線学会で得られた情報として、米国ではすでに 1,000 から 10 万症例までデータを使用している企業が出てきている。特に、資金力のある企業あるいはベンチャーが参入しているので、開発のスピードが速いという印象を受けた。（委員）
- Arterys という企業では事後学習で年に 4 回プログラムを更新していた。他にも、CAD ではないが、年に数回性能を自動更新する機器が出てきている。その場合、特に不具合がなければ、性能変化の都度 FDA への届出は行わず、現在のところ届出が必要なケース（不具合）も起きていないとのことだった。日本も早急に対応しなければ、取り残される心配がある。（委員）
- 不具合となるケースを医師の主観で全て決定可能なのかという大きな問題もある。不具合通知機能はあっても良いが、それに拘り過ぎると大変になる。米国では、既に関連学会がメーカーにフィードバックし、不具合だけでなく、臨床上役に立つ AI が評価するシステムまでできている。現在、FDA はそれと緩やかに連携している程度である。海外との競争に負けないために、国内規制は厳しすぎない方が良くと思われるが、「関連学会との連携が望ましい」等の文章はあった方が良くと思われる。（委員）
- 対象となる支援システム等の性能変化に対しては、メーカーが規定した範囲内におさまるかどうかを検証することでその品質を確保するしかないと思われる。本評価指標には「テストデータに関してはメーカーが用意したものを用いる必要があり」と記載してあるが、最低でも性能変化前の性能は確保できていることが重要であること

から、メーカーが使用した承認申請時のデータによる検証により保証できる等を明確に打ち出すのはどうか？その点を明確にすれば、本評価指標により様々な変更による性能変化も取り扱えると思われる。（委員）

- ・ 定期的なチェックに関して、対象機器によっては「毎回」必要な機器もあれば、「定期的」で良い機器もあると思われるので、使用開始前の性能評価は「毎回」としなくても良いと思われる。（委員）
- ・ 今回の対象機器は、従来の医療機器と異なり、市販後ある程度経過すると性能が変化するため、性能変化を順次検証・処理していくイメージで作業するとより実情に即した評価指標やガイダンスが作成できると思われる。（委員）
- ・ 5/(2)/2項「性能変化に伴う品質確保」で、オンサイトでユーザが事後学習できる機器に対して要求されている「支援システム等メーカーの専門家が医療機関毎に常駐して当該支援システム等の検証・管理を行う」の記載は、開発側の視点では非現実的だと思われる。（開発 WG 事務局）→前回の委員会で、オンサイトで性能変化させる場合、リスクを考慮するとメーカーと同等のマネジメントシステムが必要との意見を受けて該当の記載となった。ただし、この記述はメーカー側で他に対策を講じることを妨げる意図ではない。（事務局）→5/(2)/2 項には、その他に 2 つ、「その検証・管理が可能となるメーカーと同等のマネジメントシステムを医療機関に導入する」、「メーカーの定期的な監査実施を義務付ける等、当該支援システム等のリスクを勘案した対策を講じることが求められる。」と記載されているが、メーカーと同等のマネジメントを病院で行うのは難しいと思われる。メーカーの定期的監査にしても、それをあえて実行するメーカーは少ないと思われる。（開発 WG 事務局）→「義務付ける」とあるとメーカーにとっては厳しいと思われる。（委員）→記載内容に関しては、この後の合同会議で開発 WG 委員からも意見を収集し、それを参考に対応したい。（事務局）

5-4. 次回会議までの流れ

次回審査 WG 会議の開催日を考慮し、本日より一週間を目処として評価指標案（資料 4）に関するコメントを収集し、事務局で内容の修正を行う。この作業により精査した評価指標案を第 6 回審査 WG 会議以前にメール配布する予定である。

5-5. 審査 WG 第 6 回会議について

日時：平成 29 年 1 月 19 日（金） 14:00 - 17:00

場所：TKP 品川カンファレンスセンター ANNEX

以上

次世代医療機器・再生医療等製品評価指標作成事業
医療機器等に関する開発ガイドライン（手引き）策定事業
平成 29 年度 人工知能分野 審査・開発 WG 第 2 回合同会議
議事概要

1. 開催日時：平成 29 年 12 月 11 日（月）17:13 -18:04
2. 開催場所：AP 東京八重洲通り 13 階 A 会議室
（東京都中央区京橋 1-10-7 KPP 八重洲ビル）
3. 出席者（委員：座長以下五十音順・敬称略）
<審査 WG>
座長：橋爪 誠（九州大学）
委員：佐々木 毅（東京大学）、清水 昭伸（東京農工大学）、田中 聖人（京都第二赤十字病院）、中田 典生（東京慈恵会医科大学）、藤田 広志（岐阜大学）、待鳥 詔洋（国立国際医療研究センター国府台病院）、横井 英人（香川大学）
厚生労働省（厚労省）：青柳 ゆみ子
医薬品医療機器総合機構（PMDA）：岡崎 譲、菅原 明彦、加藤 健太郎、桐山 瑤子、古森 亜矢、宮崎 生子、今川 邦樹
事務局（国立医薬品食品衛生研究所）：薮島 由二、中岡 竜介、加藤 玲子
<開発 WG>
座長：横井 英人（香川大学/審査 WG 委員も兼任）
委員：片倉 由紀子（富士フイルム株式会社）、清水 昭伸（東京農工大学/審査 WG 委員も兼任）、野波 徹緒（オリンパス株式会社）、藤田 広志（岐阜大学/審査 WG 委員も兼任）、古川 浩（東芝メディカルシステムズ株式会社）、宮崎 靖（株式会社日立製作所）、森 健策（名古屋大学/Skype にて遠隔参加）、諸岡 直樹（株式会社島津製作所）
経済産業省：尾畑英格
オブザーバー：扇谷 悟（AMED）、浅沼 直樹（AMED）、岩田 倫明（AMED）、中里 俊章（PMDA）、宮本 大誠（PMDA）、田中 弘（日本画像医療システム工業会）、吉田 眞司（オリンパス株式会社）
事務局（産業技術総合研究所）：鎮西 清行、坂無 英徳、白澤 江身子
4. 配布資料
資料 1：議事次第（案）
資料 2：出席者名簿（審査 WG）
資料 3：出席者名簿（開発 WG）
資料 4：人工知能技術を利用した画像診断支援システム等に関する評価指標修正案
資料 5：開発 WG における検討状況

5. 議事内容

5-1. 開催挨拶及び資料確認

開会にあたり、審査 WG 座長、開発 WG 座長より挨拶があった。事務局より本合同会議の議長として開発 WG 座長を推薦し、異議なく全会一致で承認された。また、審査 WG 事務局が進行役を務めることとなり、配布資料の確認が行われた。

5-2. 評価指標案に関する説明・討議（資料 4）

評価指標案修正案について審査 WG 事務局から下記の事項が説明された。

- ・ 1 項「はじめに」の緒言を追加した。
- ・ 2 項「本評価指標の対象」における記載の詳細化を行った。
- ・ 3 項「用語の定義・説明」において、機械学習と深層学習を分けて定義した。
- ・ 4 項「本評価指標の位置づけ」に変わりはないが、審査 WG での議論において、再学習するシステム以外にも対象となる支援システム等があり得るとの指摘があり、修正する予定である。
- ・ 5 項「人工知能を利用した医療機器における問題点と解決策について」では、(1)ブラックボックスとしての性質及び(2)性能変化に関して、問題点と具体的な解決策について記載した。
- ・ 6 項「評価に対して留意すべき事項」における記載の詳細化を行った。転移学習に関する注釈を追加し、安全性、品質及び性能評価に関して、市販前と市販後に分けた上で、市販後に関しては実際に性能変化を実現する方法毎に必要な留意点を整理した。
- ・ 6 項に(4)項「リスクマネジメント」を追加し、リスク対策を追記した。
- ・ 7 項「機器の特性に応じた追加留意事項（案）」を追加し、病理診断の場合や、動画像を扱う診断支援システムの場合における留意点について記載した。教師なしデータによる学習が可能な機器も評価指針の対象としたいとする意見があったことから追記することとした。

その後、評価指標の内容について討議した。主な質疑応答は以下の通りである。

- ・ 既存の「コンピュータ診断支援装置に関する評価指標」ではデータの信頼性保証が厳しく求められていたが、本評価指標では如何か？（開発 WG 委員）→データに求められる要件に関しては大まかな方向性を記載している。性能評価にはテスト用データが最重要であるため、その妥当性を示すことが要求されると思われる。対象となる支援システムの性能等に応じた説明が必要になると考えられるため、審査 WG で具体的な要求事項を規定することは困難であると考えている。（審査 WG 事務局）
- ・ 6/(4)項「生じた不具合を漏れなく製造販売業者及び規制当局に通知する機能を必要に応じて付与することが望ましい」との記述があるが、不具合の定義が不明確な状況では実現が難しい。この要求事項が評価指標案に残された理由を示してほしい。（開発 WG 委員）→不具合の定義については審査 WG 会議でもかなり議論になり、より明確にする必要性は承知している。なお、審査 WG 事務局が想定していた不具合とは、当初規定した性能範囲が達成されず、明らかに誤った診断結果を示すケースであり、医師の判断と異なった場合には必ずその事象を通知すべきという趣旨ではない。（審査 WG 事務局）→開発ガイドラインの中で、具体的な事例を示す事は可能である。ある性能指標に異常が生じたときに取るべき具体的な対策をモデルケースとして示す等の記載方法が考えられる。例えば、特定の条件で不正解が発生しやすい傾向が明らかになった場合に適切な対応を取る等を記載できるのではないかと考えている。（開発 WG 事務局）→100%正しい判断をすることはあり得ないため、一度規定外の判断をしたから「不具合」とすることは難しい。どのようなケースを不具合と判断し、その対応策の例示を開発側で検討していただくのも良いと思っている。（審査 WG 事務局）
- ・ 通知と回収報告との関係も気になっている。6/(4)項が残っているのは回収になるまでの報告みたいなイメージなのか？（開発 WG 委員）→この文言は、ネットワークを介して人工知能が一括でメーカーによって管理されている場合に、不具合の連絡がユーザに示されるシステムを備えていると有益であろうと考えて記載した。医療

機器の回収報告等へと繋げることは想定していない。(審査 WG 事務局) → 不具合の定義や不具合事例の報告頻度等について検討する必要があると思われる。既存の医療機器においては不具合報告に関するガイドラインはないことから、そこまで踏み込んだ評価指標になるのかと不安に感じている。(開発 WG 委員) → 不具合の定義や報告機能については、開発 WG とも情報共有をしながら、引きつづき審査 WG で検討する。(審査 WG 事務局)

- ・ 4 項「本評価指標の位置付け」が今の表現のままでは、対象となる性能変化の要因が「再学習」のみになってしまうと、審査 WG 会議において指摘された。ニューラルネットワークの軽微な変更により性能変化する支援システム等も対象とすることが確認されたので、表現を修正する。(審査 WG 事務局)
- ・ 6/(2) 項「品目の原理、機能に関する基本的事項」にて、「学習アルゴリズム」や「診断用ニューラルネットワーク構造」が、「開発設計時のプログラム概要」に含まれることが分かるように表現を修正した。(審査 WG 事務局)
- ・ 6/(3)/2) 項①の「メーカーがバージョンアップ毎の段階的な性能変化を意図する場合の手続き」に、「対象となる支援システム等の性能変化において必要となる手続きについては、PMDA と事前に相談することが望ましい」という記載があるが、現時点でその手続きの対象範囲は定められていないのか。(開発 WG 委員) → まだ議論はまとまっておらず、どのような記載内容とするかは未定である。(審査 WG 事務局) → 細かすぎる等の理由で評価指標から削除した部分があれば、開発ガイドラインにおいて言及することも可能である。(開発 WG 事務局) → 人工知能技術を利用した画像診断支援システムが承認された前例がないため、評価指標には考え方を記載しておき、具体的な薬事手続きについては別の通知等で補足することになるのではないかと考えている。(PMDA)
- ・ 6/(3)/2) 項②「性能評価に必要となるテストデータ」は、学習の前後での性能比較を可能とするため、メーカーが「承認申請時に使用したもの」との説明を追記した。(審査 WG 事務局) → 「申請時に使用したテストデータ」では、初期状態に戻すことを想定しているように解釈できる。バージョンアップしている場合等は、最初の状態にまで戻す必要はないので、「そのバージョンを評価したデータ」のような表現の方が良いのではないか。(開発 WG 委員) → 表現については検討をお願いしたい。(開発 WG 委員)
- ・ 6/(4)/2) 項「内部データを利用したキャリブレーション機能」の記載を「当該システム等の内部データを利用したキャリブレーション機能」という表現に修正する。(審査 WG 事務局)
- ・ 6/(4)/2) 項「ユーザトレーニング等の適切な対策」において、「関連学会等との連携が望ましい」等の文章を追記した方が良いという意見が審査 WG 委員より出されている。(審査 WG 事務局)
- ・ 深層学習は実行する度に結果が変わるうえ、認識率が上下に大きく変動しながら確率的に学習が進むので、その点も留意した上で学習過程のデータを示すことも記載してはどうか。(開発 WG 委員) → 追記を検討する。(審査 WG 事務局)
- ・ 5/(2)/2) 項「性能変化に伴う品質確保」において、オンサイト学習により性能が下回った場合の届出あるいは監査を義務づけるのは過剰要求となることが懸念されたことから、この記述に関する議論が審査 WG で行われた。性能が低下した場合の対処法も含め、メーカーに所属する委員からこの記述に対する意見はないか。(審査 WG 委員) → メーカーの専門家が医療機関毎に常駐することも、医療機関がメーカーと同等のマネジメントシステムを導入することも難しいと思われる。リモートメンテナンス等による障害の有無の検査や、定期メンテナンスはおそらく行うことができる。(開発 WG 委員)

- ・ 6/(4)/2)項「リスクマネジメントの対策について」において、ユーザトレーニングを関連学会と連携して行うべきとの意見が審査 WG より出されているが、メーカとして実施可能か。(審査 WG 委員)→海外ではオンサイト学習できるプログラムがオープンソースでリリースされており、国内に急速に入ってくる可能性がある現状で、このように自主規制すると海外との競争において不利になる恐れがある。一方、クオリティーの確保という観点では、5/(2)/2)項に「臨床判断結果については専門知識を有する医師による確認」との表現があるように、学会がオーソライズした専門家によるユーザトレーニングの実施が大事だと明記することで、一定の歯止めをかけられるのではないか。(開発 WG 事務局)→学会がユーザトレーニングを行うのは理想的ではあるが、体制ができるまでには時間がかかるため、現時点での対応は無理と思われる。(開発 WG 委員)→関連学会等と連携してユーザトレーニング等の適切な対策を講じていく等のイメージを持っていた。(厚労省)→本評価指標の対象支援システムの開発において、海外では速いスピードで開発が行われていることから、国内規制が足かせにならず、様々な課題に対応できなければならない。両者をうまく噛み合わせながら討議していく必要があるが、企業側からみて、データの担保や性能変化について最低守るべき事項や、開発を妨げられると思われる事項があれば意見が欲しい。(審査 WG 委員)→海外メーカを排除したいという意図ではないが、このような新しい医療機器を上手くスピード感を保って活用するためには学会との連携は重要と思われる。機器の対象が広すぎてその連携をスムーズに進めることができないケースもあると思われるが、その場合には医療安全を担保しつつ、新たな医療機器の迅速な普及を図るための適切な連携先を、審査 WG や PMDA 等が判断する必要があると思われる。学会との連携に関する文章を本文に記載しておいた方が良いのではないか。(審査 WG 委員)→性能評価に対する理解や、使用方法を含め関連学会と連携して作り上げていくことが理想ではあるが、連携方法は今後の課題である。審査 WG もしくは開発 WG のどちらで扱うかも含めて検討していく必要がある。(開発 WG 委員)

5-3. ガイドライン案の進捗状況に関する報告・討議

資料 5 に従って、開発 WG 事務局より開発 WG 委員会における検討の進捗状況について下記の説明があった。

- ・ 資料 5 は、これまで開発 WG 会議における配付資料や、委員の発言よりキーワードを収集し、カテゴリー分けしたものである。各話題の右側に記載されている数字は、11 月 17 日版の評価指標に関連すると思われる章番号を記載している。
- ・ 今後、このリストを参考にして、ガイドラインのまとめ方を検討する。
座長から以下の補足コメントがあった。
- ・ 開発ガイドラインでは、評価指標で言及し切れなかった部分や、実際の開発の直接的な手法等を中心に記載していくことをイメージしている。
- ・ 既存の CAD 開発ガイドラインをアップデートする形式とするか、もしくは評価指標に対応付けるか等の議論する予定である。
- ・ 評価指標に関する議論で挙げられた課題等を反映しながら検討を進める。

5-4. その他

審査および開発 WG 委員から 12 月 20 日までに評価指標に対するコメントを収集することとした。なお、開発 WG 委員からのコメントについては開発 WG 事務局が取りまとめて審査 WG 事務局へ送付することになった。

以上

次世代医療機器・再生医療等製品評価指標作成事業
人工知能分野 審査 WG
平成 29 年度 第 6 回会議 議事概要(案)

1. 開催日時：平成 30 年 1 月 19 日（金）14:00 -17:02
2. 開催場所：TKP 品川カンファレンスセンターANNEX カンファレンスルーム 5
（東京都港区高輪 3-13-1 TAKANAWA COURT 3F）
3. 出席者（委員：座長以下五十音順・敬称略）
座長：橋爪 誠（九州大学）
委員：佐々木 毅（東京大学）、清水 昭伸（東京農工大学）、田中 聖人（京都第二赤十字病院）、中田 典生（東京慈恵会医科大学）、藤田 広志（岐阜大学）、待鳥 詔洋（国立国際医療研究センター国府台病院）、横井 英人（香川大学）
厚生労働省（厚労省）：青柳 ゆみ子
医薬品医療機器総合機構（PMDA）：岡崎 譲、菅原 明彦、加藤 健太郎、桐山 瑤子、
谷城 博幸、今川 邦樹、若尾 りか
事務局（国立医薬品食品衛生研究所）：中岡 竜介、加藤 玲子
4. 配布資料
資料 1：第 6 回審査 WG 会議議事次第（案）
資料 2：出席者名簿
資料 3-1：第 5 回審査 WG 会議議事概要（案）
資料 3-2：第 2 回審査・開発 WG 合同会議議事概要（案）
資料 4：人工知能技術を利用した画像診断支援システム等に関する評価指標修正案
追加資料 今後の予定等

5. 議事内容

5-1. 開催挨拶及び資料確認

座長挨拶後、事務局から配布資料の確認を行った。

5-2. 評価指標案の修正に関する事務局説明及び討議

第 5 回会議における議論及び会議以降に寄せられたコメントを基に作成された評価指標案修正版に関して、事務局から以下の説明があった。

2 項：用語の定義・説明

<事務局説明>

本文の内容を理解しやすくするために、「本評価指標の対象」より前に「用語の定義・説明」を移動した。また、2/(2)項「人工知能」の定義が、画像診断支援システム用に特化した内容となっていることが指摘されたことから、一般的な人工知能の定義となるように修正した。

- ・ 2/(2)項「人工知能」は、「脳能力」という表現ではなく、「人の高度な知能によって行われた推論・学習」と修正したほうが良い。（委員）→そのように修正する。（事務局）

3 項：本評価指標の対象

＜事務局説明＞

次世代評価指標の対象が「医療機器」に該当するものであることを明確にする必要があるとのコメントを受けて、対象となる支援システム等の基本要件を追加した。また、エンドユーザによる事後学習が可能な支援システム等の場合、品質管理をメーカーが行うことが難しいこと等、その医療機器該当性判断が困難となることが指摘されたため、当該支援システム等に関して考慮すべき点は別添として取りまとめた。

- ・ 対象機器が「医用画像診断支援装置」と記載されている箇所がある。意図的に書き分けているのか。(委員) → 1 項では意図的に書き分けている部分も存在するが、それ以外は全て「医用画像診断支援システム」へ修正する。(事務局)
- ・ 「エンドユーザ」という用語が出てくるが、その定義が必要ではないか。(委員) → 支援システム等を実際に使用する医療従事者を意図して記載した。必要であれば定義に追加する。(事務局)
- ・ 「個体ごとに別々の性能変化が可能となる」支援システム等の具体例が不明瞭である。(委員) → 納入された病院毎に、独立した事後学習が可能な支援システム等を想定している。この場合、全く異なった性能を持つ支援システム等が多数生まれてしまう可能性があり、結果的に現行の薬機法で求められている品目の同一性が保てなくなる。そのような製品を医療機器として規制するのは難しい可能性があるという指摘があった一方、現実にはその開発が進んでいる。ここまでの議論を全て取り込むために、次世代評価指標の対象はあくまでも医療機器であることも考慮して、本文は明らかに医療機器と判断できる支援システム等のみを対象とする形に修正し、それ以外の内容は別添として取りまとめることとした。(事務局) → 本来、この審査 WG で議論すべき支援システム等は別添の対象となったシステム等ではなかったのか。(委員) → 当初はその予定だったが、説明した通り、別添の対象とした支援システム等に関しては、その医療機器該当性も含め解決すべき課題が多々存在することが明らかとなってきた。本 WG は来年度も継続するため、引き続きその点の整理もお願いしたい。(事務局) → 今の説明だと、「個体ごと」ではなく、「施設ごと」と修正したほうがその趣旨が分かりやすくなる。(委員) → そのように修正する。(事務局)
- ・ 製品に関しては、バージョン毎に一律の性能を要求するという理解で良いか。(委員) → その理解で差し支えない。なお、その要求を満たす場合であれば、性能がリアルタイムで変化する製品も本文の対象となるように修正している。(事務局)
- ・ 画質改善や放射線治療計画にディープラーニングを用いた製品が開発されているが、これらの製品も本評価指標案の対象になるのか。(委員) → 低線量で撮像した粗い画像を改善する等の技術は対象外になると思うが、今後検討が必要だと思われる。(委員) → 米国では、画像改善に関する機器が認可されている。(委員) → 画像改善を目的とした機器は教師付き学習データを集めやすいことから、本 WG の対象機器よりも先に上市化される可能性が高い。そのため、別の目的に応用することを念頭に、画質改善を強引且つ意図的に行う機器が開発されることを危惧している。(委員) → 今回、本評価指標案の対象外となっても、医療機器であれば PMDA との個別相談が必要になる可能性があることから、過度に心配する必要はないと考える。(厚労省) → 治療計画に人工知能を用いる場合、基本的な技術は「検出」であり、その機能は本評価指標案の対象になると考えられる。(委員) → ご指摘の通りだが、将来、人工知能を利用した外科治療支援システムも開発されると考えられるので、治療計画に用いる機器を本評価指標案の対象に含めるか明確にする必要がある。(委員) → 外科用ナビゲーションシステムにおいては検出、診断、治療がセッ

トになっており、切り離すことはできない。(委員)→治療支援システムの中に診断支援機能が含まれていると考えられることから、「診断支援機能を含むシステム」も本評価指標案の対象とするよう、記載を修正するのはどうか。(委員)→今回の対象は、単なる支援システム等ではなく、ディープラーニングのように性能が変化する人工知能技術を利用した支援システム等であることを明確にしておく必要がある。(委員)→修正を事務局に依頼する。(委員)

4 項：本評価指標の位置付け

＜事務局説明＞

人工知能分野の専門家以外には、「ニューラルネットワーク構造の本質的でない一部分の選択肢を限定することでその最適化も容易に行えるようになった。」という記載の意味がわかりにくいとのコメントがあった。また、本評価指標案における「国内外の関連ガイドライン等」が存在する場合、それらの明示が必要であることも指摘されたことから、これらのコメントへの対応を委員に依頼する。

- ・ 「一部分の選択肢を限定することでその最適化を行う」との記述は、特徴マップと呼ばれるマップの数、ネットワーク層の数、畳み込みネットワークのサイズ等のパラメータを調整することで、バリデーションデータに対する性能を最大化するという意味である。(委員)→寄せられたコメントの趣旨は、事後学習の中身との違いを明確にする必要があることと理解している。(事務局)→事後学習でも、ネットワーク結線の重み決めの最適化に相当することも行われるので、その違いを明確にした文章にする必要がある。(委員)→委員に文案作成を依頼する。(委員)
- ・ 国内外の関連ガイドラインとして例示できるものはないか。(事務局)→FDA が人工知能を利用した医療機器に関するガイダンスを出している。(委員)→医療機器プログラムの承認申請に関するガイダンスも、考慮すべきガイドライン等として例示することができる。(厚労省、PMDA)
- ・ 「メーカーがネットワークを介して支援システム等により臨床で得られた結果を回収するとともに自動的に事後学習データとして人工知能に実装する」との記載は意味が分かりにくい。(委員)→臨床で得られた結果が良い場合、自動的にそのデータを回収し、事後学習に使用するという意味か。(委員)→人工知能がメーカーに存在し、臨床で得られたデータが直ちにアップロードされ自動的に事後学習に使用されることで、リアルタイムに変化した性能が実装されるものを想定した。(事務局)→単純に自動的にデータを集めて自律的に学習、性能変化する支援システム等を想定していたのであれば、前半部分は削除可能である。(委員)

5 項：人工知能技術を利用した支援システム等の評価における留意点と解決の方向性について

- ・ 口語調の箇所は修正が必要である。また、5/(1)項「ブラックボックスとしての性質」に記載された「想定外の挙動に気づくことができるような医師のみが使用するなど使用者に一定の要件を求める方法等」の記載は意味がわかりにくい。(委員)
- ・ 5/(2)/2)項「性能変化に伴う品質確保」に記載された「事後学習機能を防ぐための機能」は、「事後学習を防ぐための機能」と修正したほうが良い。また、「使用者」は「エンドユーザ」と同義なので、何れかに統一すべきである。(委員)
- ・ 米国では、性能変化における予期しない挙動に関する市場からのフィードバックを積極的に収集することが製品の上市化を促進できるという考えに基づき、専門家の集団からなる学会と FDA とが連携してフィードバック収集を目的とした組織構築

を進めている。将来、日本も規制当局と臨床現場とが連携した同様の組織構築が求められると考えられるので、そのための提言を本評価指標案に記載することはできないか。(委員)→人工知能技術にはブラックボックスの部分があり、事前に全てを把握することができないため、市販後の運用を手厚くすることを想定しながら本WGの議論に参加している。米国のように、市販後の取り扱いを調整することで上市化を促進する方策もあるということも理解している。(PMDA)→先行している米国とある程度共通の概念がないと、審査時に困らないか。(委員)→今の段階で、そのような組織の構築を要求するのは難しい。(PMDA)→評価指標はあくまでも次世代医療機器を対象に、その評価方法や評価項目を示すために作成している。医療機器評価のために望まれる詳細な枠組みを評価指標に記載するのは難しいが、あとがきとして、今後、評価において望ましい学会との連携及びその具体的な体制を提案することは可能だと思われる。(PMDA)→厚労省やPMDAの取り扱いにもよるが、報告書の中にそのような枠組み構築を提言することもできると思われる。(委員)→対象となる機器の承認後、学会と協力しなければならないケースはこれまでも存在する。特に、リスクの高い機器の場合、審査の段階から学会と協力体制を構築して市販後の取り扱いを話し合うことがある。本評価指標案の対象となる支援システム等でも同様の取り扱いとなる可能性が考えられるので、委員が所属する学会に協力を依頼する可能性はある。但し、その際は、個別製品毎の対応となる。(PMDA)→別の医療機器においては、学会と規制とが協議して、承認後の教育体制や指導体制を学会と連携して構築することを企業に求めているケースがある。報告書の中に、そのような提言を入れ込んではどうか。(委員)→人工知能技術を利用した製品が他と決定的に異なる点として、継続学習以外に転移学習による製品開発が可能であることが挙げられる。他の目的で開発された製品に、事後学習を行うことで異なる目的の製品開発が可能であるため、製品開発が指数関数的に進むことが予想される。これに対応する目的で、米国FDAは学会と連携した組織構築が必要と判断したと思われる。日本でもそのような組織構築を進めるべきだと思われるが、この分野の技術進展は著しいため、米国でも組織の詳細は定まっていないのが実情である。そのような現状を十分認識した上で、今後の議論を進めてはどうか。(委員)

6 項：評価に対して留意すべき事項

- ・ 6/(1)項「基本的考え方」2 行目に「人工知能の変化における学習形式、性能変化方法、市販後に使用する学習データの種類等に応じて規定方法や…」と記載しているが、その規定方法が何を指しているのか分からない。また、前後の繋がりが不明瞭である。(委員)→この前に「性能に影響する要素、効果のできる範囲や限界等を規定する必要がある」と記載した内容を受けた文章であり、「規定方法」はそこで記載されている要素、範囲及び限界等を規定するための方法を示している。しかしながら、ご指摘の文章は内容が不明瞭なため、その意図が分かるように修正する。(事務局)
- ・ 6/(2)/1項「原理」では、診断アルゴリズムがブラックボックス化するという点と、事後学習による変化とは分けて議論すべきである。ブラックボックス化したアルゴリズムに対して要求すべきことはネットワーク構造やプログラムの概要であり、事後学習に対して要求すべきことは使用したデータのみ明らかにすれば良い。原案には、ブラックボックスで、且つ事後学習も行う場合の要求事項のみが記載されているため、事後学習を行わないブラックボックス化した製品、あるいはアルゴリズムはブラックボックス化していないが事後学習により性能変化する製品に対しての要求事項が示されていない。(委員)→現在の記載内容は、ディープラーニングを利用した支援システム等を念頭においたものであるが、現状に即していないという

ご指摘か。(事務局)→議論はディープラーニングが主になっているようだが、問題の本質はブラックボックス化と事後学習である。事後学習を含めるということは、ディープラーニングのみでなく他の機械学習も対象になるということである。ディープラーニングを念頭において記載した内容であることは理解できるが、他の機械学習も対象となりうる上、現実に出てくる製品には様々な機械学習を適用してくることが考えられるので、ディープラーニングを前面に押し出しすぎるのは好ましくない。ブラックボックス化と事後学習とを分けて、明らかにすべき要求事項を記載すれば、結果的にディープラーニングを使用した場合に要求すべき事項も網羅できる。(委員)→通常の医療機器プログラムではアルゴリズムに関する詳細な記載が求められるが、今回対象となる支援システム等ではそれが難しいため、その代替法が記載されていると思われる。事後学習も従来の医療機器にはない概念であるため、その説明が記載されていると理解する。(委員)→6/(2)/1項「原理」の中に学習及び事後学習が記載されているが、これは分けて 6/(2)/2項「学習」の項に記載すべきである。また、事後学習の時点で、学習アルゴリズムまで変更される可能性があるため、それに対応できるよう修正を行うべきだと考える。(委員)→ブラックボックス性に対して記載すべき内容、学習に対する懸念に対して明らかにすべき内容が整理されれば良いと考える。(PMDA)→6/(2)/1項「原理」で明らかにすべき項目は診断、あるいは検出のためのアルゴリズムであり、学習のアルゴリズムではない。一方、6/(2)/2項「学習」で、学習アルゴリズムを明らかにすることを要求すべきである。加えて、市販前後で学習アルゴリズムを変更する可能性も考慮した内容を追加すべきである。(委員)

- ・バリデーションという用語の使い方を整理すべきである。バリデーションとテストを混同されてしまうと、適切な性能テストが実施されない可能性がある。(委員)→機械学習のデータには学習用、バリデーション用及びテスト用が存在することにも注意すべきである。(委員)
- ・6/(2)/2項「学習」に示された「学習前に加工している場合にはその処理方法」は、**Data Augmentation** で人工的にシミュレーションした場合や、画像を加工した場合とは異なるのか。(委員)→**Data Augmentation** 等以外による加工が存在する可能性を考慮して記載した。(事務局)→学習に使用することが前提であれば、**Data Augmentation** で全ての加工が含まれるので、その記載は不要である。(委員)→匿名化のような加工は考慮しなくても良いか。(委員)→学習データとするためのプロセスではないため、匿名化に関する記載は別途必要である。(委員)→6/(2)/3項に個人情報の取り扱いに関する記述を追加している。(事務局)→6/(2)/2項で記載された加工の意味が「匿名化」を意味しているのであれば、こちらで対応できているので、やはり記載は不要と考える。(委員)→ご指摘に従って、6/(2)/2項の記載は削除する。(事務局)
- ・6/(3)/1項「市販前」の中に市販後に関する事項が混入しているように思える。(委員)→市販前の段階で、市販後に起こりうる性能変化範囲を規定しておくことが必要となることを明記した。(事務局)→同一製品と判断されるためには、承認申請時にその性能変化範囲を規定しておく必要がある。(委員)
- ・6/(3)/1項「市販前」の2ポツ目に「以下に示す検証用テストデータに関する詳細な情報」とあるが、情報の範囲が不明瞭である。(委員)→3ポツ目以降に示された検証用テストデータに関する情報を全て示すことを要求しているわけではなく、製品の目的、仕様に応じて項目を適宜選択して明示することを想定した記載だったが、ご指摘の通り、必要な情報の範囲が分かりにくい内容となっている。(事務局)→メーカーにとっては重要な情報なので、表現を再度検討してほしい。(委員)
- ・6/(3)/1項「市販前」の4ポツ目に記載されたテストデータの「頑健性」の意図は何

か。(委員)→支援システム等の性能を担保するためには、テストで規定された範囲のアウトプットが出ることが重要であるとの前提で議論が進んでいたことから、テストデータを第三者が勝手に改変することができないよう適切に保護することを要求事項として記載した。(事務局)→開発メーカがテストデータを知りすぎると、そのデータを学習データとする可能性もある。そのため、テストデータはメーカに対しても秘匿する必要があるという記載にしてはどうか。(委員)→その場合、メーカがテストデータを用意できなくなるので、その準備方法を本評価指標案で示すことが必要になる。(事務局)→テストデータは事後学習の場合も継続して評価対象となるので、この議論は重要である。メーカがいつでもテストデータにアクセスできる状態にあると、テストデータの信頼性に疑義が生じることになるため、それを防ぐ手段を講じておく必要がある。ただ、現在、テストデータを準備するために使用可能な大規模データベースは存在しない。現実的には企業が準備せざるを得ないことから、適切なテストデータの管理方法を設定する必要がある。(委員)→これまでの議論を個人的に理解した限りでは、テストデータの適切な選択や管理方法は開発 WG で議論したほうが良いように思われる。(事務局)→懸念事項は、テストデータを学習データやバリデーションデータとして使用されてしまい、性能テストの結果が恣意的にコントロールされてしまうことだと理解した。データの選択にバイアスが生じないよう、メーカが関与しない方法で画像を収集してテストデータとすることが望ましいと思うが、そのような方法は現実的ではない。承認審査ではその収集方法を明確に示してもらうことが現実的対策で、その説明から妥当性を判断するのが、経験上、限界だと感じている。本評価指標案には、選択バイアスがかからないように配慮してテストデータを準備することを要求するのが無難ではないか。(PMDA)→修正案には、データの種類、入手元及び妥当性を詳細に示すことが要求されているので、これで良いのではないか。(委員)→上市后、製品は消費者という第三者によるテストを経て取捨選択され、性能が悪いものは残らなくなるので、テストデータについて厳密な規定を行う必要はないと思われる。(委員)→しかしながら、上市后、実際に使用された結果、性能が劣っていたということが判明すると臨床現場は混乱してしまうので、ある程度のテストデータ管理は必要になる。そこは審査でも確認されると思うが、秘匿化を求めるか、メーカが示した管理方法を信頼するか等、製品に応じて温度差が生じてしまう可能性はある。(委員)→テストデータに関しては、頑健性よりもバイアスがかかっていないことを求めることが重要という理解で良いか。(事務局)→普通の審査でもバイアスがかかっているので、「バイアスがかからないように配慮する」という表現が現実的な対応だと思われる。(PMDA)

- 6/(3)/1項「市販前」の最終段落にある「十分な説明が可能となる質と数のデータ」との記述は「質と量のデータ」と修正したほうが良い。(委員)
- 6/(3)/2項「市販後」では、エンドユーザが受け取る情報を、具体的に「その動作や性能の詳細」と記載したほうが良い。また、「データに偏りが生じないように」との記載があるが、ここまでの議論にもあったように、データの偏りを完全に排除することは難しいことから「データの偏り等にも配慮しながら」と修正すべきである。(委員)
- 6/(3)/2項「市販後」に「事後学習により検証なしにデータが追加され変化した」との記載があるが、ここまでの議論から検証なしに性能を変化させるものは認めないと考えていた。この「検証なし」は、どういう意味のものを指しているのか。(委員)→ここは、臨床現場で使用した結果をリアルタイムに反映させて性能を変化させるものを想定した記載となっている。(事務局)→リアルタイムの場合でも、その変化した性能を評価した結果、規定範囲外になれば性能変化を止めるのではない

か。(委員)→そのようなケースに備えて、リアルタイムの場合は事後テスト機能やキャリブレーション機能を要求するという趣旨の文章としたつもりである。(事務局)→「検証なし」を削除すれば、その趣旨は明確になる。(委員)→ご提案の通り、削除する。(事務局)

- 6/(4)項「リスクマネジメント」に、リスクと考えられる事象として「人工知能が想定以上に誤った情報を提示すること」と「誤ったデータの学習による性能低下」が挙げられているが、その両者の違いが分からない。(委員)→故意に不適切なデータで事後学習をさせた結果、あるいは誤った臨床結果を紐付けしたデータを事後学習させた結果、支援システム等の性能が低下することを想定した。(事務局)→性能が低下する、あるいは誤った結果を提示する原因は多数存在する。その中に、誤ったデータによる学習もあると思うが、その原因のみが強調されることに違和感がある。前者のみを記載すれば十分だと思われる。(委員)
- 6/(3)/2)②項「臨床に伴った自動的な学習により連続的又は相当程度高い頻度でのバージョンアップを意図する場合」で、メーカーに定期的な性能検証を実施することを要求しているが、詳細な検証方法まで記載するのか。例えば、あるシステムに対するテストを実施し、その結果により規定された性能範囲内に収まっているかを評価するのであれば問題ない。一方、一つのシステムに対して 10 通りの事後学習を行い、それらを全てテストした後に最も優れていたものを選択してアップデートした場合、テストに用いたデータはテストデータではなく学習データとなる。後者のケースを防止するために、検証方法に関する詳細な記載が必要だと考える(委員)→市販後の性能評価方法に関しては、種々の方法が設定可能である。定期的な検証もあれば、自動検知機能を備えておきユーザからの連絡に応じて適宜対応を行う方法も想定される。(委員)→基本的には定期的な性能検証は必要だと考えている。しかしながら、バッチ式でアップデートする支援システム等の場合、ユーザから誤った判断が行われたという情報があった場合のみ、再度メーカーが検証を実施することになると考えられる。一方、クラウドタイプの人工知能を使用しリアルタイムに性能が変化する支援システム等については、(4)項「リスクマネジメント」に定期的な性能検証を行う間隔を定めるよう要求事項を記載している。詳細な記載は必要だと思われるが、性能検証及びそのために使用するテストデータに関する考え方は基本的に市販前と変わらないことを示しているため、市販後でも性能検証時には適切なテストデータを使用することを要求していることになる。ご指摘の点に関しては、むしろ、テストデータの取り扱いの中に記載すべきではないかと思われる。(事務局)→市販前、市販後という問題ではないため、テストに関する部分に適切な記載を追加すべきである。(委員)
- 6/(4)項「リスクマネジメント」に「適用外使用」という表現があるが、「適応外使用」が正しい用語である。(委員)→ご指摘の通り、修正する。(事務局)
- 6/(4)/2)項「対策について」の 1 ポツ目の記載内容に関しては、(2)/2)項「学習」の記載内容が参考となるので、その項を参照するように記載するのが適切だと思われる。(委員)→そのように追記する。(事務局)
- 6/(4)/2)項 2 ポツ目の「内部データを利用したキャリブレーション機能」はどういう意味か。(委員)→支援システム等の中に性能キャリブレーションのためのテストデータを準備しておいて、使用前点検のように、当該システムを立ち上げるたびにそのデータを利用して性能検証を行う機能を想定していた。前回の会議で、このような機能を備えていけば、少なくとも承認申請時の性能は担保可能との提案が委員からあったため、ここに記載した。(事務局)→自己テストとの違いが不明瞭である。(委員)→ご指摘の通り、「キャリブレーション」を削除し、「自己テスト」のみとする。(事務局)

- ・ 6/(4)/2)項 3 ポツ目に「人工知能を初期化」とあるが、これは市販時の状態に人工知能を初期化することを求めているのか。(委員)→初期化に関しては、市販時の性能に戻すことを意図した記載である。明確にするため、「人工知能を市販時の状態に初期化」と修正する。(事務局)

7 項：機器の特性に応じた追加留意事項（案）

<事務局説明>

関連する委員が作成した原案を事務局で体裁を整え修正した項である。基本的には原案を全て反映した内容としているが、補足の必要性も含めて確認を委員に依頼する。

- ・ 7/(1)項「事後学習による性能変化が診断支援においてハイリスクとなる場合」で例に挙げられた病理診断で、求められる性能は「検出率」ではなく「鑑別率」ではないのか。(委員)→ご指摘の通りなので、適切な用語を検討する。(委員)
- ・ 原案では、病理専門医の関与があれば自律的に学習を行う支援システム等は認めるという内容に理解される。(委員)→この文章の趣旨は、性能変化のための事後学習には病理専門医の関与が必須であることと理解している。(委員)→性能変化には病理専門医の関与が必要という文章に変更したほうが理解しやすい。(委員)→学習のみでなく、その使用においても病理専門医の関与が必要ではないか。(厚労省)→ご指摘の通り、AI 懇談会の中で、最終的な診断は医師が行うことになっている。今、問題になっているのは、学習に使うデータが病理専門医の関与なしに使用することである。(委員)

別添「エンドユーザ等が性能変化に関与可能な人工知能技術を利用した医用画像診断支援システム等に関する基本的考え方（案）」

- ・ 別添の最終的な取り扱いは、本評価指標案が通知化される際に厚労省、あるいは PMDA により判断されるため、報告書にのみ掲載される可能性もある。(事務局、委員)
- ・ 別添で取り上げるべき点は「施設ごとの事後学習」だと思われるので、事後学習を誰が行うかは明示する必要がない。なお、実施者にかかわらず全てに適用できる議論であれば、「エンドユーザ」は「エンドユーザ等」と修正すべきだと思われる。(委員)→これまでの議論で、事後学習の実施者としてメーカーとエンドユーザの両者を想定する必要があると考えていた。特に、エンドユーザが事後学習を実施すると、メーカーが実施する検証よりも質が劣ってしまう懸念が指摘されていたことから、原案ではそれらを分けて記載している。(事務局)→使用現場で事後学習する支援システム等において求められる要件が、この部分に記載されているという理解で良いか。(委員)→その理解で差し支えない。前回の会議で、メーカーが関与せずに臨床現場で事後学習させることが可能なシステムの場合は、臨床現場にメーカーと同等のリスクマネジメントシステム等を設けるよう要求するべきという意見があったことから、その内容を反映させた記載としている。(事務局)→規定範囲内に収まる性能を持つ支援システム等を認めるのであれば、臨床現場で事後学習可能なシステムの場合でも、その性能範囲に収まることを前提に、追加の要求事項を満たせば認めるという内容にしたほうが良い。(委員)
- ・ 3/(1)/1)項「エンドユーザ等による事後学習が可能な支援システム等の場合」に「一定の品質確保が困難となる」との記載があるが、ここの「品質」は性能変化以外が対象という理解で良いか。(委員)→このケースでは、性能変化に伴い、承認時の品質を確保することが難しくなるということだと思われる。(委員)→「品質」には性能が規定範囲から逸脱するリスクがあることも想定されていると思われる。ユ

一ザのみが事後学習を行う場合、そのリスクがさらに高くなるという趣旨の記載が、その後に追加されているという理解で良いか。(委員)→その理解で差し支えない。

(事務局)→今の記載内容だと、そのリスク以外の問題が新しく起こるように理解されるので、修正が必要だと思われる。(先生)

- 施設毎の事後学習が可能な支援システム等は、その結果、性能がバラついてしまうので本文とは別に取りまとめたという理解している。仮に、規定された範囲内であっても、その性能がバラついてしまうのも問題にならないか。(委員)→予め規定された性能範囲内に収まれば問題ないと思われる。(委員)→その場合、施設毎に異なる性能となっても問題ないと理解される。(委員)→本評価指標案の対象となる支援システム等では、性能が規定範囲内に収まることが大前提であり、それを満たす限りその支援システム等の使用は可能である。施設毎に事後学習が可能となる支援システム等の場合、事後学習方法の結果、性能が偏ってしまい規定範囲を逸脱した支援システム等が臨床現場で継続して使い続けられる懸念があることから、本評価指標案の対象から除いた。しかしながら、これまで、このような支援システム等を医療機器として認められる範疇を会議で議論してきたことから、その内容を別添としてまとめた。(事務局)
- 3/(1)/1項で「事後学習責任者」と「事後学習データの妥当性を示すことができる責任者」を分けているが、何か意図があるのか。(委員)→事後学習データの臨床的妥当性を示すことは専門性を有した医師のみが可能と思われるが、支援システム等の原理及び事後学習方法を熟知した人物として医師の他に支援システム等を開発した技術者等が想定されたので、各々の知識を有する責任者が別々に指名される可能性を考慮した。(事務局)→ご指摘は、同一人物が2つの役割を兼任することを禁ずると誤って理解される可能性を意味したものではないか。(PMDA)→分けて記載する必要は感じない。むしろ、ここに記載された項目が責任者となる人物には必要である等を示した文章にしたほうが良いと思われる。但し、全ての項目を充足することが必須と理解されないように、書き方を変更すべきである。(委員)→ご指摘に従って、修正を検討する。(事務局)
- 3/(1)/2項「教師なしデータによる学習が可能な機器の場合」で「支援システム等がどのような形で開発されるかは不明な点も多い」と記載されているが、不明な点はこのような支援システム等が事後学習によって性能を変化させる原理やその方向だと思われる。この場合には、人が全く管理しないため、リスクが非常に高くなることを危惧しているのであれば、「どのように事後学習が進んでいくのか不明な点も多い」と修正するのが良い。(委員)

5-3. 総合討議

- 6/(2)/3項のタイトルに記載されている「セキュリティ」は非常に広い範囲を示すことから、「使用環境及び情報セキュリティ」とするほうが良い。(PMDA)
- 5項のタイトルが長すぎる一方、6項のタイトルが短くてアンバランスだと思われる。(委員)→「人工知能技術を利用した」という枕詞は全体を通して削除する。(事務局)→「特徴の整理と解決の方向性」としてはどうか。(PMDA)→検討して修正する。(事務局)
- 5/(2)/4項「責任の所在」には、性能変化に対する責任だけでなく、支援システム等全般に対する責任も言及していることから、この位置に記載するのが相応しいのか。(委員)→全般に対する責任、特に医師がその機能に依存しすぎないようにするための記載であるため、ご指摘の通り位置を変更する必要がある。(事務局)→7項に移動させてはどうか。(PMDA)→追加留意事項ではないので、この位置での記載は不適切だと思われる。(PMDA)→5/(3)項としてはどうか。(委員)→「孕んでい

る」という用語も修正する必要がある。例えば「含んでいる」と修正してはどうか。

(PMDA) → ご指摘の通り、修正する。(事務局) → 最終的には医師が主体性を持って診断することが必要であることを明確にするため、先に提案があった通り、記載箇所を 5/(3) 項とするのが良いと思われる。(委員)

- クラウド活用を含んだガイドラインは、今回初めて作成されると思われる。人工知能と関係ない部分もあるが、クラウドに関する留意点を書き込む必要はないか。(委員) → クラウドとスタンドアロンタイプの医療機器ソフトウェアとの違いとして、その経路を含めたセキュリティの問題が挙げられる。この問題については、医療機器ソフトウェアに関連したガイダンスで言及されていたと記憶している。(委員) → 当該ガイダンスを関連ガイドラインの部分に特出しして記載してはどうか。(厚労省) → 6/(2)/3 項「使用環境と情報セキュリティ」に参考ガイダンスとして明記してはどうか。(事務局) → 全体の参考ガイダンスとなるので、4 項「位置付け」の部分に記載するほうが望ましい。(PMDA) → 追加になるが、6/(2)/3 項の 3 ポツ目に記載されている「端末コンピュータがインターネットと接続」は「支援システム等がインターネットと接続」と修正したほうが良い。(委員)

5-4. 今後の流れ等

- 本日の討議を元に事務局で本評価指標案の修正版を作成し 1 月 25 日を目処にメール配布し、各委員で再度確認する。報告書へ掲載する分担報告事項の締め切りは 2018 年 2 月 5 日とする。
- 本評価指標案に対する意見募集を連携学会に依頼するか否かについては、各学会代表者の委員を通じて学会に確認を取り、事務局へ連絡する。
- 報告書には「評価指標案」として掲載されるが、最終的な評価指標(通知)となる前に、厚労省を通じてパブリックコメントを募集するため、今回、報告書の査読依頼を見送っても、学会の意見を反映させる機会はある。(厚労省)
- パブリックコメントの募集期間が判明次第、本 WG 関係者へ連絡する。(事務局)
- 本評価指標案別添は、本年度の報告書に掲載されるが、通知化される際に添付されるか否かは現時点で明言できない。いずれにしても施設毎に事後学習が可能となる支援システム等に関しては来年度の WG で再検討する。(事務局)
- 報告書は年度明け早々、国立衛研医療機器部の HP に掲載予定である。(事務局)
- 学会の理事会等で説明する際の資料として、パブリックコメント募集を含めた今後のスケジュールに関する資料が欲しい。(委員) → ロードマップに準じる資料を作成しメールで WG 関係者へ送付する。(事務局)
- 報告書としてまとめる前に、本評価指標案を各学会の理事長に送る必要はないのか。(委員) → 元々、理事長を介して委員の選出をお願いした関係上、参加委員から理事長に提出し、学会内で審議していただくのが良いと考えている。(事務局) → 基本的にこれ以上、書き換えをしないということか。(委員) → 個人的に以前関わった WG は、いずれも学会と連携していない。一方、別の WG では学会推薦専門家が委員として参加されていたこともあり、学会に査読を依頼した結果、1-2 件のコメントが寄せられた経験がある。本評価指標案の対象は大変活発な分野であることから、より多くの意見が得られる可能性はある。(事務局) → 期限を切っても意見を募るべきではないか。(委員) → 原則、パブリックコメント時に学会から意見を募る方向が良いのではないか。(委員) → 学会によっては多数のコメントが寄せられると思われるため、一般向けのパブリックコメントで意見を募ることに同意する。(委員) → 学会毎に可能な対応を依頼することで良いと思われる。寄せられたコメントに関しては、年度を跨いだ場合でも、来年度見直しの際に有効活用していくことを学会側に併せて伝えていただきたい。(事務局) → 来年度以降継続して審

議していくことを前提にしたコメント募集であれば、各学会からの積極的な本 WG への関与が期待できる。(委員)

- 本 WG は来年度も継続する。想定している来年度の作業は以下の通りである。
 - 世界的な開発・規制動向に変化がある場合、それに対応するための改定作業。
 - 評価指標案別添に記載した内容の再検討。
 - 国内外の人工知能技術の医療分野全般に対する応用の現状調査。
 - 上記の討議を踏まえた上で、評価指標が必要と考えられる新規対象分野選定及び評価指標案作成。
- 厚労省全体でも AI 関連の事業に力を入れている。この分野は急速に進歩しており、FDA も様々な判断を示しつつあることから、継続的に調査・議論を進めていく必要があると考えている。必要に応じて WG の拡大も視野に入れた検討をお願いしたい。(厚労省)

以上