

「アルゴリズムが「創る」ガバナンス・倫理」  
日本経営倫理学会 ガバナンス研究部会 永井郁敏

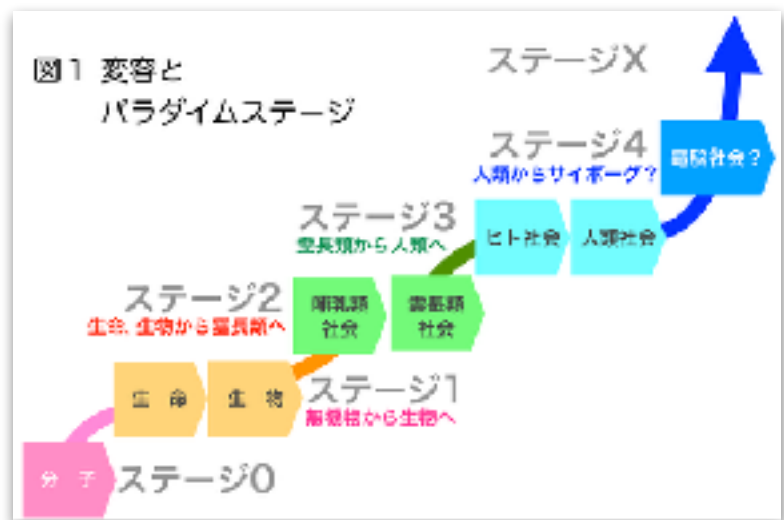
## 1. はじめに

現代社会は自動化の真っ只中にある。その自動化社会は、様々なプログラムをはじめとするアルゴリズムで出来上がっているといても過言ではありません。このアルゴリズムには、意図的なバイアスと無意識なバイアスがもたらす「偏り」が存在します。そしていまこのアルゴリズムは「技術先行、ガバナンス後追い」により、多くの不公正や差別を生み出す中、アルゴリズムの公正さを如何に実現するかが、問われはじめています。

技術力がもたらす新たなパラダイムに向け、バイアスを認識し自動化システムに潜在するリスクを検討することから、アルゴリズムへの公正さとガバナンスの必要を考えます。

## 2. 令和に起こる社会変革

AI（人工知能）の時代、そのAIからのアウトプットを決定するものの1つにアルゴリズムがある。そしてこのアルゴリズムは、人工知能やその1つ手前の言わばオートメーションの延長線上にあるRPAの結果を決定すると言っても過言ではありません（但しここでこのロボットは、人型やハードウェア稼働のロボットではなく、ソフトウェアロボットを前提とする）



この人工知能やロボット技術、そして生命科学の進歩は、いずれも高速コンピュータとネットワーク技術が実現するものであり、今後30年の令和を育むために欠かすことのできない技術だと考える。そしてこれらの技術が引き起こす今回のパラダイムシフトは、人が過去に経験したことのない量と速度で、変容することになる（図1）

## 3. いま起きていること（人間に代わってシステムが自動選択する時代）

インターネットのサービスは、本人の購買履歴から次に買うべきものを予測し、事前に勧めるリコメンド機能や、まとめ買いによる更なる値引きを勧める機能など、本人のライフスタイルや個人の特性に合った商品を、システムが対象者個人の情報を分析し、自動的に判断して勧めて来る。同様に、航空便のチケットやホテル料金、そしてプロ野球の観戦料金も、コンビニエンスストアの仕入れの決定も、航空機の自動操縦も航空・交通管制に至る多くのサービスや公共システムなど、これらは人間の意思決定を要することのない自動化されたシステムで稼働しています。

そして最も急速な技術革新の先鋒がAI（人工知能）であり、特に機械学習とディープラーニングの2つが自動化に拍車をかけている。そしてこれら多様な自動化社会を実現するソフトウェアが、アルゴリズム（algorithm）なのです。

#### 4. アルゴリズムとは

アルゴリズムを図式化したものをフローチャートと呼んでいます。またIT用語辞典の説明では「ある特定の問題を解く手順を、単純な計算や操作の組み合わせとして明確に定義したもの（後略）」としている。



図2  
アーキテクチャー/アルゴリズム/アプリケーション

同様なカテゴリに、アーキテクチャとアプリケーションがある。(図2)は全体設計のアーキテクチャ、手順を示すアルゴリズム、個別の作業を実行するアプリケーションとの関係を表した図式ですが、ドイツのインダストリー2.0をコンセプトとした我が国のSociety5.0(超スマート社会)は、アルゴリズムとしての表記は少ないものの、この図2の3つの概念を外すことはできません。



左図は内閣府の作成による Society5.0の概念図の1つだ。破線部分(筆者加筆)ではAIにより必要な情報が必要な時に提供される社会と、ロボット技術などにより人間の可能性が広がる社会の実現を目指している。

しかし現在は、その一歩手前であるRPAやAIを稼働させるために求められる「デジタル・トランスフォーメーション(DX)の時期」だと筆者は考えている。

DXをひとことで言えば「アナログからデジタルへ」の変容だと言える。このデジタル化は、コンピュータ(algorithm)が判断できる情報にしなければならないことから、欠かすことができない。コンピュータが判断できる様にするためには、外部環境や、物の状態やプロセス、そして人の五感値など、あらゆるものを計量化・数値化し、データとして取り込むことが求められるからです。

#### 5. 自動化の実際

ITとRPAそしてAIの一側面を、自動化を切り口に表したものが次ページの図式(図3)ですが、以降では実際の自動化を可能にしている3領域を、具体的にみて行きます。

##### 5-1. RPA(事務処理のソフトウェア)

Excelで集計したデータを基幹システムに自動的に入力したり、ネットバンキングへ自動でアクセスして入金情報を収集したり、事務処理を自動化できるソフトウェアが「RPA(Robotic Process Automation: ロボティック・プロセス・オートメーション)」です。

このRPAは、ユーザインタフェースで稼働し人間の作業をそっくり再現できるため、バーチャルのBPO（ビジネス・プロセス・アウトソーシング）を可能にすることから、日本でも2020年以降大幅な利用促進が期待されています。未だまだ！

しかし既に海外では「RPAガバナンス」をコンセプトに、SOX対応や監査対応、そしてコンプライアンスとしてRPAの適用範囲が広がっており、RPAガバ

ナンスは、RPAの導入目標達成を担保するための「経営システム」として位置付けられています。残念なことに我が国では未だその認識は少なく、コスト削減を実現するツールとしての認識に留まっているのが現状ですが、米国の4大会計ファームでは、RPA+Blockchainを用いた内部統制システムを稼働させ、RPAが自動生成したプログラムコードに、監査の対象がシフトしています。これは自動生成されたプログラムコードへの編集ができないことが、SOXへの適応を可能にしている。

現在その特性から、RPAは定型業務に適していますが、IBMの人工知能であるIBM Watson®との連携で、人の介在が求められる非定型業務にもその範囲を広げています。事務作業が自動化され人の介在が無い分、不正も改ざんも起こしにくく、対象データをブロックチェーンに格納することで、データへの改ざん記録が全て白日の下に置かれ、その信頼性は一層高まることとなります。5年先の2025年には内部統制、会計監査などの監査関連業務は、既に人の手を離れていると考えられる。

## 5-2. スマートコントラクト（ブロックチェーン）

スマートコントラクトとは、元々は1990年代に法学者・暗号学者であるNick Szaboが提唱したとされており、暗号化通貨技術の基盤であるブロックチェーンに実装されている技術で、直訳すれば賢い（スマート）契約（コントラクト）となる。契約をスムーズに行うことを目的とし、契約の定義から実行までを、第三者が介すことなく、すべて自動で行えるところに特徴があり、自動販売機のプロセスが代表例として用いられている。

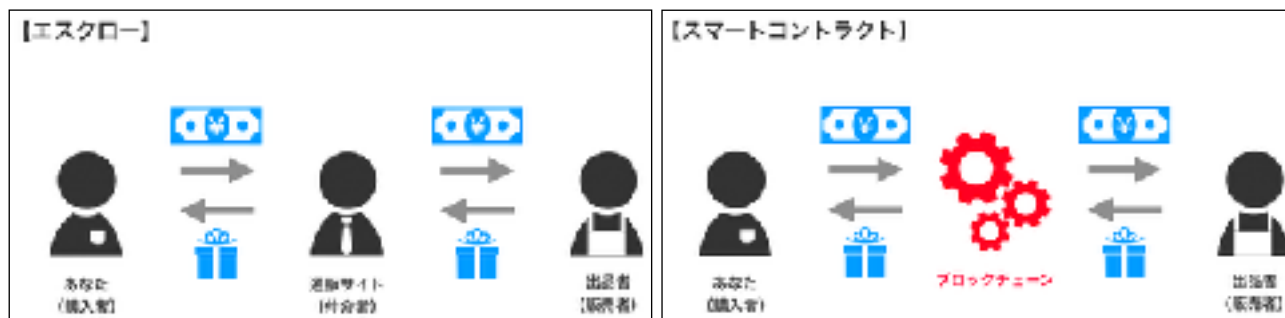
- ① 利用者が契約条件を満たす額のお金を投入し特定のボタンを押すと、その商品（飲料など）を利用者に提供する取引契約として予めプログラミングする
- ② 利用者が契約条件を満たす額のお金を入れて特定の商品ボタンを押す
- ③ その商品が利用者に提供される
- ④ お金が商品の販売者のものになる（決済される）

この取引契約のプロセスのうち、②～④を自動化した仕組みがスマートコントラクトで、取引契約をスマートコントラクトで実行することで、取引に欠かせない「相手を信用する



コスト」の大幅な削減を可能にする。ブロックチェーン上の取引は公開されている上に、改ざんできないため、安全性と透明性が向上するわけです。そして何より、スマートコントラクトでは、取引の仲介者として第三者を挟む必要がない。これまでは銀行等の金融機関や信販会社（決済代行会社）等が、取引の安全性を保証（下図左：エスクロー）していたが、ブロックチェーンによって「信任された第三者」の中抜きが可能になる（下図右）

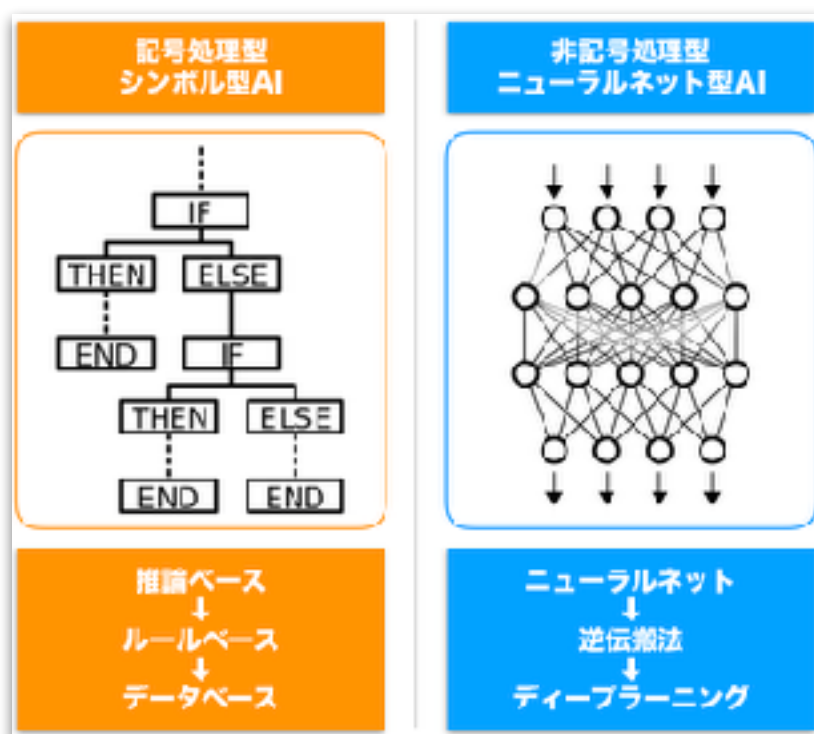
[出典] Udemy, Inc.



### 5-3. 人工知能 (AI) のアルゴリズム

AIの定義は、工学や科学そして哲学など、各分野の専門分野からの異なった視点で言及されるため、未だ確かな定義は存在していないと筆者は考えているため、今回はその一部について言及するに止めたい。

このアルゴリズムは、現在特にAIに象徴されている部分が多いが、その分類方法の考え方として、IBM Watson®やGoogle検索等に代表されるシンボルによるAI（記号処理型）と、いわゆる機械学習、ディープラーニングと言われるニューラルネット型のAI（非記号処理型）に大別できます（下図：三宅洋一郎等に筆者加筆作成）



シンボル型は数式を用いた推論型 (if…then~) であり、さまざまな数式を用いて、データの処理や問題の解決に向けたアクションを起こし、与えられたデータをもとに、正確な解決へと導くのが特徴となる。

一方のニューラルネット型は、「パターン認識」もしくは「機械学習」と呼ばれる学習、分析、認識などを自分で行う「ニューラルネットワーク」と呼ばれるアルゴリズムが使われ、顔認識や自動翻訳など、与えられた情報の特徴

を、人が判断条件として与えられることなく、AI自身がその特徴点を見出すことを可能にしている。

学習を行ったAI（図4左側）は、その学習データを基に入力した図柄や文字音声などの与えられた情報（図4右上）を、パターン化された学習済みデータ（図4左側）と比較し判断を行うアルゴリズム（図4右下）により結果を出力する。

現在、最新の機械学習の主流は、記号処理型から非記号処理型へ移りつつありますが、このどちらにも、開発者が意図するしないに関わらず、アルゴリズムがもたらす結果を左右する偏りの存在がある。以下では、この偏りがもたらすリスクと課題を考えていきます。

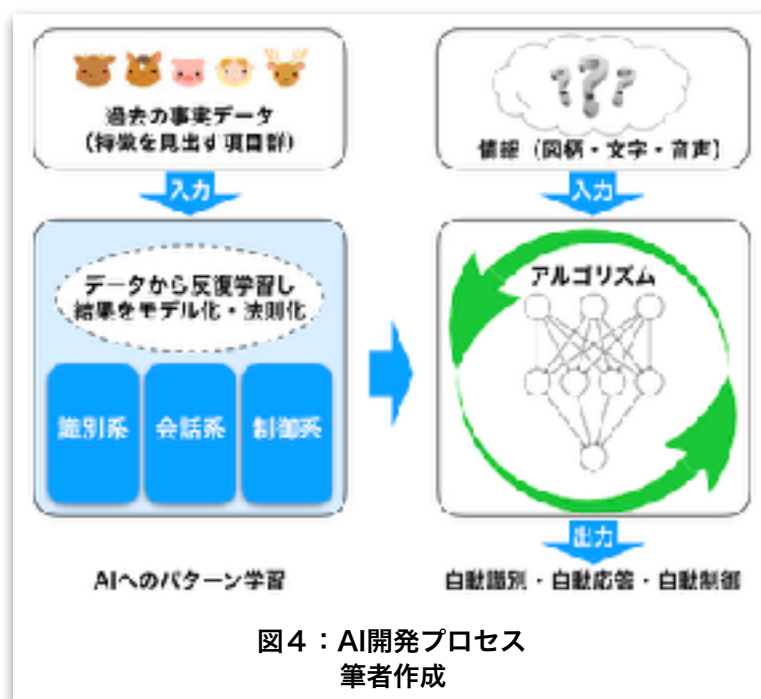


図4：AI開発プロセス  
筆者作成

## 5. 何が問題になるのか？（見えない「リスク」：意図的な操作と無意識の偏り）

では、ここまでに挙げたRPAやスマートコントラクト、そしてAIのアルゴリズムの何が問題になるのでしょうか？。コンピュータである人工知能は「機械」であり、アルゴリズムが受け取ったデータを無言で処理し続けるため、人はその結果を公平・公正であると決めつけがちになる。アルゴリズム（数学）で動くアプリケーションがデータを基に経済を動かすとしても、基になるアプリケーションは人間の選択で作られているわけで、開発者である作り手の先入観、固定観念、誤解といった偏見（バイアス）が、これらアルゴリズムの中に潜在することになる。

偏りを意図的に生み出すことが、どこまで倫理的に問題があるかは一概には言えないが、例えばインターネットの広告では、人間の弱い心理を突くことで、広告主はその目的を達成させている（ex.利用者がAGAだとわかると、育毛剤やウィック、専門クリニックの広告が頻繁に現れてくる）。これらの操作は、利用者個人の「認知バイアス」をターゲットに、その思考の癖とも言える部分に意図的に訴求しています。

人は自分の「思考の癖」である認知バイアスをほとんど意識していないため、このようなことが起こるわけで、このバイアスが、アルゴリズムを開発する技術者に、偏りの意識がないまま取り込まれ、見えない偏りとして、AIのアウトプットに影響をもたらすことになる。つまり用いるデータの中にも、アルゴリズムの判断プロセスの中にも、そしてアウトプット結果を検討するヒトの潜在意識にも、様々なバイアスが潜んでおり、誤った結果を生み出している。これらのバイアスは、RPAを制御するアルゴリズムやAIを開発するアルゴリズムにも、ブロックチェーンのスマートコントラクトにも存在する。

そしてこのバイアスには、開発者本人が気づか無いバイアス「無意識バイアス」と、何かの目的のための作為的なバイアス「意図的バイアス」がある。

AIの専門家であるニール・ジェイコブステインは、人間が持つ次のような認識傾向（無意識バイアス）を、AIやアルゴリズムを用いることで、改善したり補完することができる」と解説している。認知バイアスの種類の理解にも役立つので紹介する。

- **アンカリング**：意思決定の際、情報の一部や特定の特徴のみを過度に重視する傾向
- **可用性バイアス**：最近あったことの記憶や珍しいこと、感情を揺さぶられた記憶が影響を及ぼす結果、特定の出来事が起きる可能性を過大に評価してしまう傾向
- **確証バイアス**：自分の先入観を肯定するような情報を探したり、重視したり、覚えていたり、あるいは肯定するように情報を解釈したりする傾向
- **フレーミング**：情報を提示する人物や提示方法により、同じ情報から全く異なった結論が導かれること
- **楽観バイアス**：過度に楽観的になり、好ましい情報や結果を過大評価してしまう傾向
- **計画誤差**：利益を過大評価し、コストや計画達成までの時間を過小評価してしまう傾向
- **埋没費用／損失回避バイアス**：何らかを諦めるというデメリットの方が、手に入れるというメリットよりも大きく感じられる傾向

他方「あなたを支配し社会を破壊する、AI・ビッグデータの罠」の著者キャシー・オニールは、意図的なバイアスとして、

- 結果として、恵まれた人はより恵まれるようになり、不利な状況にある人はより不利になる。アルゴリズムは結果を予測しているのではない、そうなるよう仕組んでいる。
- データサイエンスの危機は特に見えにくい、知らない間に作られたシステムが、私たちの社会を根底から崩している。
- ネットを使う全ての人、ターゲット型広告を駆使する業者に狙われる可能性が高く、アルゴリズムはソーシャルメディアや検索履歴など膨大なデータから顧客の痛みに訴求する。情報や教育で弱い立場の人が最も犠牲になりやすい。

と、意図的なバイアスへの警鐘を鳴らしている。

## 6. まとめ（理想社会への提言）

右図はIAA（国際内部監査協会）がAI監査のフレームワークとして図式化したもので、そのコアにはAI戦略／ガバナンス／人的要因が位置付けられている（右図）。

このフレームワークでは、AIにおける内部監査の役割は「人工知能が短期、中期または長期の価値を創造する組織の能力に、否定的または肯定的な影響を与える程度を評価し、理解し、伝達するのに役立つこと」としています。

また赤十字国際委員会（ICRC）のホワイト



ペーパー「武力闘争におけるAIと機械学習（本年6月）」においても、その人道的アプローチとして、多方面の視点による注意点を挙げている。これは人工知能の開発課題にある「フレーム問題」として、アルゴリズムが「意味やコンセプトを理解できない」ことから、多方面の視点によるチェックの必要性が求められている。

同じ様な状況下でも、AIは人間とは全く違う判断をすることがあるため、AIによるアルゴリズムでの処理は、特に敵対的状況下での弱さが指摘されています。

このような状況を踏まえ、我が国においてアルゴリズムに関する広義のAI監査システム（組織や制度も含めた）を、早急に確率することが急務であると考えています。アルゴリズムを導入し、明確で公正な配分を行うためにも、このアルゴリズムに対するガバナンスと監査、そして倫理を合わせ、早急に検討・対応する必要があります。

しかしそのためには、ITとAIの専門家、そして内部監査人はもとより、倫理や哲学者、科学者、宗教学者などなど、凡そ関連するプレーヤーによる「当該エコシステム」を早期に稼働するべく、AI監査プラットフォームの構築が必要だと考えています。

つまり、このAI監査PFは単一の組織や団体といった従来の枠組みで思考するのではなく、個々の業界とその上位の組織といった全体的な視点でAI人工知能を開発&利用する、或いは影響を与える側と受ける側等、およそ関係する全てのプレーヤーで検討する必要があるからです。

何故ならば、このアルゴリズムがもたらす世界（つまりAIがもたらす世界）は、組織や業界といった単一の系で閉じることができないからです。関係する様々な技術や科学、大学や学会の視点やスキルがハイブリッドに関係する点が、今までの単一技術とは明らかに異なっています。そしてこれは正にAIが必要とする知能（視点）であり、AI的視点に他ならない。

冒頭に図1で示したステージ4の新たな技術が牽引する社会変革（エクスポネンシャルな変化）は、余りにも長くダーウィンの進化論的な直線的变化の中にいた我々の感覚では、中々理解しづらいところではある。

ことが起こってから対応すれば良いという考えや、同業の状況を横並びで対応するといった日本的慣習が未だ数多く存在しますが、昨日までの前提が変化し、もはやそこに意味や価値を失ってしまっただけからでは全てが遅すぎる。

旋風を読む感覚と、それを見える化（オープン化）し、共有し、協働するリーダーシップが火急に必要なところである。

以上

## 参考文献：

1. 経営倫理#89；BERCの実践研究レポート，永井郁敏，経営倫理実践研究センター，2018.1
2. コンピューター・アーキテクチャー；ベン・ケーシー，
3. 人工知能のための哲学塾；三宅洋一郎，ピー・エヌ・エヌ新社，2016.8
4. 人工知能のための哲学塾東洋哲学編；三宅洋一郎，ピー・エヌ・エヌ新社，2018.4
5. 脳がつくる倫理；パトリシア・S・チャーチランド，化学同人，2013.8
6. あなたを支配し、社会を破綻するAI・ビッグデータの罠；キャシー・オニール，インターシフト，2018.7
7. 飛躍する方法；サリム・イスマイル，日経BP，2015.8
8. ファスト&スロー（上下巻）；ダニエル・カーネマン，早川書房，2012.12.25
9. 予測マシンの世紀（AIが駆動する新たな経済）；アジェイ・アグラワル，早川書房，2019.2
10. 教養としてのテクノロジー；伊藤穰一&アンドレー・ウール，NHK出版，2018.3
11. 武力紛争における人工知能と機械学習～人間中心のアプローチ～；国際赤十字委員会（ICRC），2019.6
12. 日本経営倫理学会研究発表大会，統一論題資料；前野隆司，2019.6
13. AIと倫理の議論はなぜ噛み合わないのか～法律家の視点から～；林 秀弥，名古屋大学大学院法学研究科教授，2018.
14. アルゴリズムがつくる「公正さ」には、差別を助長する危険性が潜んでいる；伊藤穰一，WIRED(US)，2019.5
15. 人工知能の応用における倫理やガバナンス面の課題；Joich Ito，2019.5
16. AIのバイアス問題はなぜ解決しないのか？；Karen Hao，MITTR，2019.2
17. Artificial Intelligence: Privacy Promise or Peril?；BRENDA LEONG，FUTURE OF PRIVACY FORUM，2019.2
18. GLOBAL PERSPECTIVES AND INSIGHTS(IIA)，Artificial Intelligence – Considerations for the Profession of Internal Auditing，2017
19. UNFAIRNESS BY ALGORITHM: DISTILLING THE HARMS OF AUTOMATED DECISION-MAKING，FUTURE OF PRIVACY FORUM，December 2017.  
<https://fpf.org/wp-content/uploads/2017/12/FPF-Automated-Decision-Making-Harms-and-Mitigation-Charts.pdf>
20. Why AI is a threat to democracy—and what we can do to stop it；Karen Hao，MIT Technology Review，Feb 26，2019.  
<https://www.technologyreview.com/s/613010/why-ai-is-a-threat-to-democracyand-what-we-can-do-to-stop-it/>
21. ETHICAL OS(Risk Mitigation Checklist)；Institute for the Future and Omidyar Network，2018
22. データ利活用とデジタルガバナンス；経済産業省，2019.3
23. AIの「失敗」責任は誰が負うべきか；Karen Hao，MIT Technology Review，2019.05
24. AI利用のルールづくり目指す、世界経済フォーラムが評議会設立へ；Will Knight，MIT Technology Forum，2019.05
25. 経済産業省DXレポート；-ITシステム「2025年の崖」の克服とDXの本格的な展開~，2018.9
26. 人工知能 3；一般社団法人人工知能学会，2019.7
27. 人工知能 4；一般社団法人人工知能学会，2019.5

以上