情報処理学会論文誌 Vol.60 No.9 1427–1432 (Sep. 2019)

招待論文

デジタルトランスフォメーションに潜むチャンスとリスク

浦本 直彦[1,a)]

受付日 2019年4月9日, 採録日 2019年6月27日

概要：人工知能や IoT などのデジタル技術を活用しながら新しい製品サービスやビジネスモデルの創出や企業風土の変革を行うことで，新しい価値の創出を目指すデジタルトランスフォメーションが現在注目を集めている．デジタルトランスフォメーションにより，新しい価値を生み出すとともにデータに基づいた経営判断を迅速に行うことが期待されている一方で，これまでに想定していなかったリスクが生じる可能性が生じている．本稿では，日本の製造業でデジタルトランスフォメーションを推進する著者の経験をふまえながら，デジタルトランスフォメーションに潜むチャンスとリスクについて考察する．

キーワード：デジタルトランスフォメーション，人工知能，機械学習，データ，信頼

Opportunities and Risks for Digital Transformation

Naohiko Uramoto[1,a)]

Received: April 9, 2019, Accepted: June 27, 2019

Abstract: Digital transformation is expected to create a new business value by creating new business models, product services, and ecosystem, and by changing organizational climate, powered by digital technology by machine learning. While executing digital transformation enables quick business decisions based on data, such emerging digital technologies may introduce risks that were not previously envisioned. This article describes opportunities and risks behind digital transformation, based on the author's experience promoting digital transformation in traditional Japanese manufacturing company.

Keywords: digital transformation, AI, machine learning, data, trust

1. はじめに - デジタルトランスフォメーションの要素

デジタルトランスフォメーションには複数の側面があり，その定義は自明ではない．2004 年にデジタルトランスフォメーションの概念を最初に提唱したスウェーデンのウメオ大学のエリック・ストルターマン教授によれば，デジタルトランスフォメーションとは「IT の浸透が，人々の生活をあらゆる面でより良い方向に変化させる」ものである．現在我々を取り巻く状況を表現するために具体的な単語をあてると，「デジタル技術を用いて新しいビジネスモデルや製品サービス，エコシステムの構築，企業や組織風土の変革を行うことで，新しい価値を創出すること」となるだろう．しかし，本稿では，筆者の製造業の現場におけるデジタルトランスフォメーションの経験から，もう少し広い意味でこの言葉をとらえたい．World Economic Forum が 2017 年に公開した “Digital Transformation Initiative Chemistry and Advanced Materials Industry” レポート [13] であげられているデジタルトランスフォメーションの要素を以下に示す．

高いレベルの効率化と生産性の向上

既存の生産プロセスやビジネスプロセスに対して，機械学習を用いた予測（需要予測，異常検知など）や，判断の自動化などの技術を適用することで，プロセスの効率化や最適化を行う．オペレーショナルエクセレンス（Operational Excellence）とも呼ばれる．

既存の製造プロセスの効率化や生産性向上は，デジタリゼーション（デジタル化）であってデジタルトランスフォメーションではないという見方もできよう．トランスフォ

[1] 三菱ケミカルホールディングス
Mitsubishi Chemical Holdings, Chiyoda, Tokyo 100–8251, Japan
[a)] uramoto.naohiko@ma.mitsubishichem-hd.co.jp

メーションの本来の意味に従えば，それは既存のプロセスの効率化ではなく，構造の変化が必要だといえるからである．しかし，生産プロセスにおける何らかの異常を機械学習を用いて検知することで，生産工程を安定化し生産性を向上させることを考えた場合，既存の演繹的手法に基づいたモデル化やシミュレーションから，データに基づく帰納的なアプローチへの変換という観点において，単なるプロセスの効率化ではないと考える．また，プロセスの可視化や自動化を通じて，プロセスの最適化につながっていく．本稿では，機械学習などのデータに基づく製造プロセスの効率化や生産性向上をデジタルトランスフォメーションの一部だとして議論を進める．

デジタル技術が加速するイノベーション

たとえば，新素材の探索にデータ科学を用いるマテリアルズインフォマティクス [16] が現在注目を集めている[*1]．膨大な実験から望みの物性を持つ化合物を探索する従来の演繹的手法から，大量のデータを用いて物性から構造を推測する帰納的手法へと研究開発のやり方が大きく変わる可能性がある．さらに，量子コンピューティングなどの新しい技術の適用も期待されている．

データ管理と知見の生成

大量のデータを分析することで得た知見をビジネス判断や生産プロセスの効率化，最適化に生かすものであり，機械学習を含む人工知能技術や IoT 技術が用いられることからも，現在大きな注目を集めている．また，製造現場においては，熟練者の技能継承が大きな課題となっており，大量のデータを知識化することで，熟練者の経験を若い世代に伝えていくことができると期待されている．

労働力への影響

デジタル技術の浸透により，従業員のスキルの変革が求められる．たとえば，データ分析をデータの意味を知る現場の技術者が行うことができれば，より深い知見が迅速に得られる可能性がある．

デジタル技術を活用した製品サービスのオファリング

B2B 企業が，直接の顧客だけではなく顧客の顧客，あるいは最終消費者の声を吸い上げるための仕組みを構築することでより高い価値を提供する．

デジタルエコシステム

プラットフォームを中心にしたビジネスモデルが従来のサプライチェーンの構造を破壊し，ビジネスの支配者が入れ替わる可能性がある．

以上，デジタルトランスフォメーションの要素について述べた．デジタルトランスフォメーションの本質は新しい価値の創出にあると考える．この新しい概念が実現されていくに従い，これまでにないリスクを考慮する必要がある．このような新しいリスクが生まれる背景として，以下の 2 つがあると仮定する．

(1) デジタル技術の急速な進化とともに，技術が複雑化し，その詳細を理解することが難しくなってきている．
たとえば，深層学習のアルゴリズムは，訓練データから学習モデルを構築するが，その中身はモデルのパラメータである数値の集合であり，人間が理解することは困難である．また，多層ニューラルネットワークを用いる深層学習においては得られた出力の根拠を理解することが難しいことが指摘されている．もちろん，これまでも非専門家には理解できない難解な技術が存在した．しかし，現在のデジタル技術は，研究開発と社会に投入されるまでの距離が近くなっている．利用者が直接目にするインタフェースは洗練されており，使いやすいけれども裏側ではどのような処理が行われているのかが分からないことが多い．また，新しいアルゴリズムが日夜開発され，GitHub などを通じてその実装を手に入れることも容易になってきた．中身の詳細を理解せずにダウンロードしたプログラムを使っていることが多い．丸山は，深層学習のような高次元のパラメータを持つ技術が人間の認知限界を超えつつあることについて議論している [18]．人間の理解がそもそも必要かという問題についてはここでは触れないが，このような技術の複雑化が新たなリスクを産む可能性がある．

(2) ソフトウェアの振舞いが人間の直感と合わなくなってきている．
デジタル技術が複雑化し人間の理解がそれに追いついていない状況下においては，それらの技術を用いたソフトウェアの振舞いが人間が予想するものと異なることによるリスクが生じる可能性がある．従来，人間と人間，人間とソフトウェア間のやりとりにおいては，(明示的な事前の合意があるかどうかは別として) お互いの振舞いに関する期待があり，信頼 (Trust) のベースとなっている．昨今のデジタル技術が提供する様々なソフトウェア (たとえば機械学習などの人工知能技術，ソーシャルネットワーク，プラットフォームなど) においては，その信頼にずれが生じている．従来技術を前提とする期待値に基づいた行動が予想しなかった結果をもたらすリスクを認識する必要がある．

次章以降では，技術の複雑さと信頼のずれを示す例として，敵対的サンプル，学習モデルの解釈性，フェイクニュース，ソフトウェアライフサイクルの変化，を取り上げる．

2. 敵対的サンプル (Adversarial Examples)

通常の機械学習では，正しいラベルが付けられた訓練データから作られた学習済みモデルに対して，新しいデータを入力し，そのデータに対するラベルを推定する．これに対し，あるサンプルデータに対し，意図的に小さなノイ

[*1] 人工知能学会誌 2019 年 5 月号では，マテリアルズインフォマティクスの特集が組まれており参考になる．

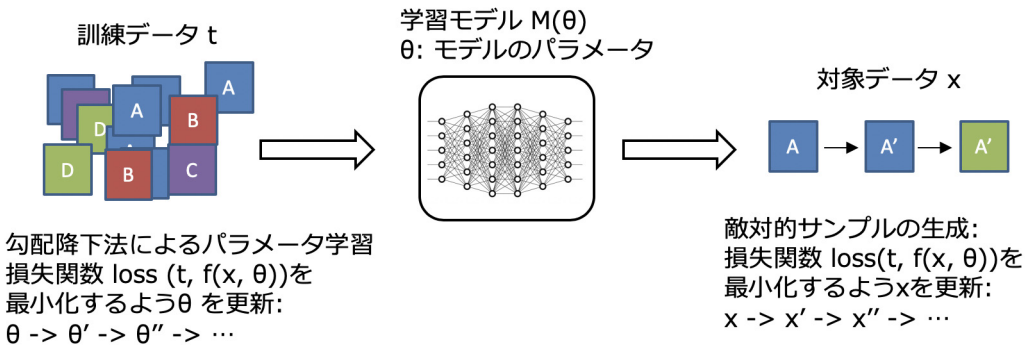

図 1 敵対的サンプルの生成
Fig. 1 Generation of adversarial examples.

ズ（摂動）を入れることで，学習モデルを欺き，誤ったラベルを出力させることが可能である．このようなデータを敵対的サンプル（Adversarial Examples）と呼ぶ．2014 年に，Szegedy らが深層ニューラルネットワークに対する敵対的サンプルの論文 [10] を発表し，大きな注目を集めた．以降，様々な敵対的サンプル生成の手法とそれに対する対策手法が提案されている[*2]．Goodfellow らは，パンダの画像に対し，摂動を加えることでその画像をテナガザルに誤認識させる例を紹介している [2]．また Eykholt らは，道路標識の画像にノイズを入れることで，停止標識を速度制限標識に誤認識させることができたことを報告している [1]．

敵対的サンプル生成の基本となる考え方は単純である．モデルの学習，つまり，ニューラルネットワークの各ノードの重みを計算するには，勾配降下法がよく用いられる．これは，対象データ x に対し，正解とモデルの出力（関数 $f(x,\theta)$，θ はモデルのパラメータ）の差異を表現する損失関数を定義し，損失が最小になるように，重み θ を更新していく．x が変更されることはない．敵対的サンプルでは，逆に，θ を固定し，損失関数を最小化しながら x がターゲットとなるラベル（上の例では，パンダではなくテナガザル）に分類されるための最小の摂動を与えて x を変化させる．この過程を図 1 に示す．

ここで，パンダの画像が，画像更新の過程でだんだんとテナガザルに変わっていくわけではないことに注意したい．人間にはパンダに見える画像をシステムがまったく別のものと誤認識することは，人間の直感には合わないため，何らかの悪用の発見の遅れやそれに起因する事故などにつながる恐れがある．

3. 機械学習モデルの解釈性（Interpretability）

技術が高度化し人間の理解が困難になってきている例の 1 つとして，機械学習における学習モデルの解釈性（Interpretability）[4], [17] がある．深層学習で用いる多層ニューラルネットワークでは，学習モデルは，ネットワークの構造と数値で表現される各ノードに対する重みの集合として定義される．また，学習モデルへの入力に対して，判別結果（たとえばラベル）が出力される．この際，なぜそのラベルが選択されたのかの根拠を知ることは困難である．

現実的な局面において，得られた結果を元に何らかの決定を行う場合，その根拠を示す必要があることが多い．たとえば，車の自動運転のプログラムに，機械学習を用いたアルゴリズムが含まれており，そのプログラムの出力結果によって事故が発生した場合，運転手，自動車メーカ，プログラムを作成した企業，データを提供した企業の誰がどのように責任を持つかを判断するためには，プログラムがなぜそのような計算を行ったかを知る必要がある．また，工場の製造工程において，機器から出力されデータを元に機器の制御を自動化する機械学習プログラムを導入した場合，安全性の見地からも，システムがなぜそのような判断を行ったかを人間が理解できないと，そのシステムがなかなか現場では受け入れられないという現実がある．

解釈性や透明性（Transparency）に関する規制やガイドライン策定も進んでいる．欧州では，2018 年 5 月に施行された一般データ保護規則（General Data Protection Regulation, GDPR）においては，個人情報を元に何らかの自動化された判断を行う際の説明責任について規定されている [3]．日本でも，内閣府が主導する人間中心の AI 社会原則検討会議において公平性，説明責任および透明性の原則が議論されている [22]．

解釈性についても，利用者にとってのリスクは，自分が理解していないシステムを使うことにあるという指摘がある [4]．入力に対して，なぜその出力が得られたかを分かりやすく説明する手法が必要となる．また，データに基づいて学習モデルを構築する機械学習では，そのデータがそもそも妥当なものかについても検討が必要である．

ここで，すべての機械学習モデルについて解釈性が必要であるわけではないことに注意したい．金銭的または安全面からビジネス上のインパクトが大きな判断を機械学習

[*2] もっとも，機械学習における敵対的手法（Adversarial Machine Learning）は，2000 年代初頭から研究されている（たとえばスパムメールの分類を誤らせる文面の生成）．

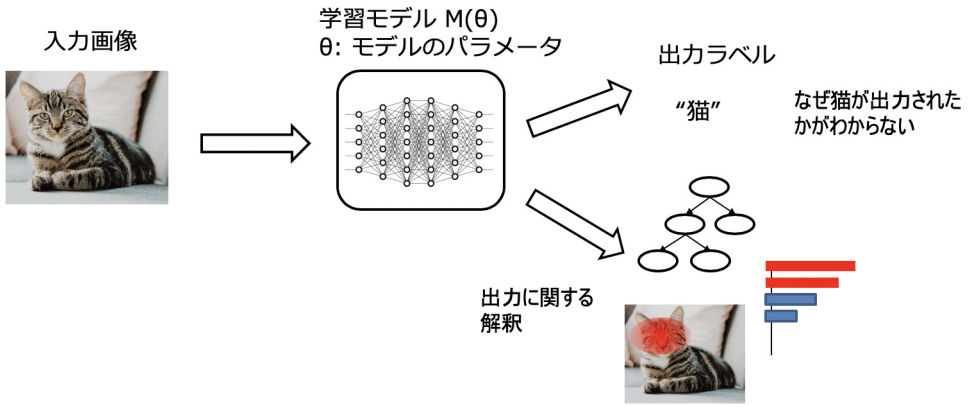

図 2 機械学習モデルの解釈
Fig. 2 Interpretation of machine learning models.

システムを活用して行う場合には解釈性が重要であるが，そうでない場合もありうる．また，機械学習システムにおいて解釈性を担保しようとすると，システムの精度やパフォーマンスが低下する可能性がある．

機械学習モデルの解釈性に関する研究は近年活発に行われており（原によるまとめ [19] が参考になる），様々な手法が提案されている．たとえば，データのどの特徴量が結果を出力するのに重要だったかを，決定木 [9] や画像領域のハイライト [8] によって提示することで理解することが可能である（**図 2**）．また，どの訓練データが出力結果に影響を与えたかを提示することで，説明を行う手法 [6] も提案されている．

また，構築された学習モデルの評価をどのように行うかについても議論が行われている．モデルの統計的有意性を評価するのに P 値という指標がよく用いられる．この値が小さいほどそのモデルは統計的に優位であると解釈でき，P 値が 0.05 より小さいことを判断の基準にすることが多い．しかし，この P 値の算出に開発者のバイアスが入る可能性が指摘されており，P 値ハッキングといわれることもある．たとえば，P 値は，サンプルデータ集合をどのようにとるかによっても変化するためバイアスが入りやすい．2015 年にアメリカ統計協会は P 値の扱い方に関する声明を出している [12]（内容の一部は日本計量生物学会によって翻訳されている（www.biometrics.gr.jp/news/all/ASA.pdf））．

4. フェイクニュース

Twitter や Facebook に代表されるソーシャル・ネットワーキング・サービス（SNS）をビジネス上の意思決定に活かす動きが広がっている．社会が急速に変化することで従来のビジネス感覚が通用しなくなり，スピード感を持って価値ある製品サービスを提供するために，社会にアンテナを張り，顧客の声を迅速に取り入れることが求められているからである．また，SNS を使った顧客への浸透をマーケティング戦略の一部とする動きも定着している．これら SNS の膨大なユーザ基盤から生みだされる情報は膨大かつ多岐にわたる．このような状況下で，SNS から発信される情報の中に，誤ったものや捏造されたものが入り込み，投票などの重大な人間の行動に影響を与えている．ここで，フェイクニュースは，広義には，特定の意図なく作成された誤情報（作成者の無知や編集システムの不備によるものなど）と，悪意を持って作成される偽情報に分けられる（偽情報のみをフェイクニュースととらえる場合もある）．フェイクニュースの議論でよく引き合いに出されるのが，2016 年のアメリカ大統領選挙やイギリスの EU 離脱の是非を問う国民投票において，虚偽の情報が有権者の選挙行動に影響を与えたとされる事件である．

もちろん SNS 上の情報でフェイクニュースに分類されるのはごく一部である．また，利用者側も信頼するサイトや友人経由の情報や，多くの他のユーザが目にして議論している情報を取り入れるなど，情報に対する信頼の判断をしているはずである．しかし，それでも騙されるのは，信頼のベースが揺らいでいるからではないかと考える．Laser ら，および Vosoughi らが 2018 年に Science 誌で発表した論文では，フェイクニュースの方が事実情報よりも拡散が早く，特に政治記事についてはその傾向が顕著であると報告している [7], [11], [14]．その理由について以下の点が指摘されている．

- 従来から，人は自分が真実だと思っている情報を裏付ける情報を信じる傾向にあることが知られている．それらの情報が SNS 上の巨大な友達ネットワークから共有されたものである場合，それを信頼してしまう傾向にある．
- 自分が好みに合わせてカスタマイズした（ニュースサイトなどの）情報源を信用する傾向にある．
- 虚偽の情報に繰り返し晒されると，自分がそれに関する知識を持っていたとしても虚偽の情報を信じてしまう傾向にある．
- 目新しいことに注目し，それを自分が知っていること

を他人に伝えたいという欲求がある．
- 有名人のスピーチに別の人間の声と口の動きを合成する動画の作成技術が進展している．

フェイクニュースへ対応するためのアプローチについては，利用者の教育や動機付け，情報が事実かどうかを判別し，フェイクニュースの拡散を防ぐ技術の開発，規制の整備などがありうる．しかし，そもそも，事実と虚偽の情報の間に明確な線を引くことはできないのかもしれない．そのような情報交換の中で，どのように現実的な信頼を確保するのかが課題である．

ここまで，代表的なデジタル技術をいくつか取り上げ，新しい技術の導入に関するリスクについて議論した．ここからは，これらの技術を用いたアプリケーションや IT システムにおけるリスクについて考えてみたい．

5. ソフトウェアライフサイクルの変化

機械学習に基づく処理を含むシステムを構築する場合には，従来とは異なる点を考慮する必要がある．**図 3** に，従来の Web アプリケーションと機械学習アルゴリズムを用いた同様のアプリケーションの構成図を示す．従来の典型的な Web アプリケーションは，バックエンドのデータベース，アプリケーションロジック，クライアントソフトウェア（Web ブラウザ）のいわゆる 3 層構造からなる．アプリケーションロジックは，ルールとしてプログラムコードの中に埋め込まれて一体化した形で扱われる．一方，機械学習アルゴリズムを用いたシステムでは，アプリケーションロジック本体は，学習モデルに埋め込まれる（プログラムコードの中に学習モデルを呼び出すコードやロジックの一部は存在する）．さらに，学習モデルは，訓練データによって構築される．ここで生じるリスクを以下にまとめる．

- 学習モデルが改変されるリスクがある．前述したように，学習モデルは人間には解釈することが難しいので，改変されたとしてもそれに気がつかない可能性がある（ニューラルネットワークのパラメータが改変されても，出力結果が変わるだけで，ソフトウェアとしては正しく機能している）．
- 学習モデルが搾取されるリスクがある．たとえば，与信システムのための学習モデルは，業務ノウハウの塊であり，それが流出することの損害は大きい．
- アプリケーションの開発と，学習モデルの構築が別の企業や組織によって行われる可能性がある．たとえば，大量の一般的な訓練データから学習したモデルと，対象領域の少量のデータを組み合わせる転移学習においては，大量のデータを持つ企業やコミュニティが提供する事前学習モデルを用いることが多い．異なる提供元のコンポーネントを組み合わせて全体のセキュリティや安全性を担保するのは自明ではない．
- 訓練データが悪意を持って改変されていたり，公正ではなかったりする可能性がある．

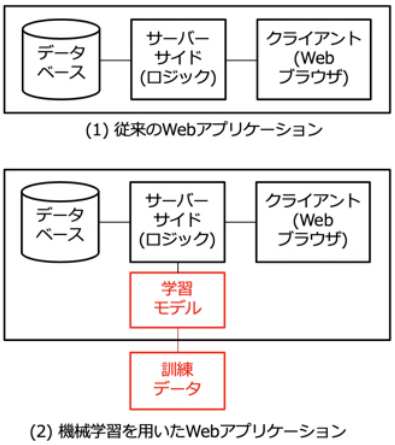

図 3 典型的な Web アプリの構成図
Fig. 3 Typical components of Web applications.

また，要件定義，設計，開発，テスト，運用，バージョンアップからなる一連のソフトウェア開発工程の指針が，これまでとは変わる可能性がある．たとえば，大規模プラントの機器から生成されるプロセスデータに機械学習を適用した異常検知システムを構築したとしよう．大規模プラントでは，数年に一度定期修繕が行われ，機器の入れ替えや構成の変更が行われる．その場合，プロセスデータの特性が以前とは異なる可能性があり，学習モデルも再構成する必要があるかもしれない．その場合，過去のデータがどのくらい有用なのかを議論する必要がある．このように機械学習を用いたシステムが主流になった場合に，これまでのソフトウェアエンジニアリングの手法が変わる可能性がある．井上らは，金融分野において，機械学習に基づくシステムに対するセキュリティ分析を行っている [15]．また，ソフトウェア科学会が機械学習工学研究会（https://sites.google.com/view/sig-mlse）を立ち上げ活動を開始しており，この分野の今後の発展が急務である．

6. データドリブンな判断を支えるために

最近は，IT 企業だけでなく，筆者が所属する伝統的な製造業においても，製造現場や経営層からデータドリブンな判断をしたいという声を聞くようになった．もちろん，データサイエンティストを雇用したり機械学習技術を用いたシステムを導入したりすることだけでデータドリブン経営が叶うわけでもなく，データトランスフォメーションの本質と同様に，組織やそこで働く社員の変革が必要であることはいうまでもない [20]．しかしこの話題を深掘りするのは本稿の範囲を超えるので，データドリブンな判断を支えるためのデータについて今一度考えたい．

第一に，取得するデータがそもそも正確なのかという問題がある．データの品質管理は昔から議論され標準化が行われてきた分野である．たとえば，ISO-5725（JIS 8402）[5]

は，化学分野などにおける実験データの測定方法や精確さの推定に関する国際標準であり，実験データの測定法やその評価については，確立された手法が存在する．しかし，製造業の現場においては，センサから送られてきたデータが大量にあるものの，現場に行ってみるとセンサが壊れていて不正確なデータを送っていたということがある．データそのもの精度や真正性について考慮することが必要である．

データそのものの精度や真正性に加えて，データが大量に収集された場合，データに偏りがないかという問題が生じる．データの公平性については，現在活発な議論 [21] がされており，公平性を計量するためのツールなどの研究開発も含めて，今後の進展が期待される．

7. おわりに

本稿では，企業や組織がデジタルトランスフォメーションを進めていく際に生じる，これまでとは異なるリスクについて考察した．複雑化する技術とそれに起因する信頼のずれに利用者が対応するのは容易ではないが，まずは，その際に生じるリスクを正しく認識，計量することが大切である．リスクを計量することができれば，それを軽減するための具体的な方策を考えたり，利益とのバランスを取りながらリスクを許容したりすることができる．デジタルトランスフォメーションは，ビジネスと企業風土の両方を変革することで大きな価値を創出する可能性を秘めている．リスクを適切にコントロールしながら社会の変革が実現していくことを期待したい．

参考文献

[1] Eykholt, K., Evtimov, I., Fernandes, E., Li, B., Rahmati, A., Xiao, C., Prakash, A., Kohno, T. and Song, D.: Robust Physical-World Attacks on Deep Learning Models (online), available from ⟨http://arxiv.org/abs/1707.08945⟩ (2017).

[2] Goodfellow, I.J., Shlens, J. and Szegedy, C.: Explaining and Harnessing Adversarial Examples, pp.1–11 (online), available from ⟨http://arxiv.org/abs/1412.6572⟩ (2014).

[3] Goodman, B. and Flaxman, S.: European Union regulations on algorithmic decision-making and a "right to explanation", pp.1–9 (online), DOI: 10.1609/aimag.v38i3.2741 (2016).

[4] Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Pedreschi, D. and Giannotti, F.: A Survey of Methods For Explaining Black Box Models, pp.1–45 (online), available from ⟨http://arxiv.org/abs/1802.01933⟩ (2018).

[5] ISO: ISO 5725-1:1994, Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions, Technical Report (1994).

[6] Koh, P.W. and Liang, P.: Understanding Black-box Predictions via Influence Functions (online), available from ⟨http://arxiv.org/abs/1703.04730⟩ (2017).

[7] Lazer, D.M.J., Baum, M.A., Benkler, Y., Berinsky, A.J., Greenhill, K.M., Menczer, F., Metzger, M.J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S.A., Sunstein, C.R., Thorson, E.A., Watts, D.J. and Zittrain, J.L.: The science of fake news, *Science*, Vol.359, No.6380, pp.1094–1096 (online), DOI: 10.1126/science.aao2998 (2018).

[8] Ribeiro, M.T. and Guestrin, C.: "Why Should I Trust You?": Explaining the Predictions of Any Classifier (online), DOI: 10.1145/2939672.2939778 (2016).

[9] Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D.: Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, *Proc. IEEE International Conference on Computer Vision*, Vol.2017-October, pp.618–626 (online), DOI: 10.1109/ICCV.2017.74 (2017).

[10] Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I. and Fergus, R.: Intriguing properties of neural networks, pp.1–10 (online), available from ⟨http://arxiv.org/abs/1312.6199⟩ (2013).

[11] Vosoughi, S., Roy, D. and Aral, S.: The spread of true and false news online, *Science*, Vol.359, No.6380, pp.1146–1151 (online), DOI: 10.1126/science.aap9559 (2018).

[12] Wasserstein, R.L. and Lazar, N.A.: The ASA's Statement on P-Values: Context, Process, and Purpose, *The American Statistician*, Vol.70, No.2, pp.129–133 (online), DOI: 10.1080/00031305.2016.1154108 (2016).

[13] World Economic Forum: Digital Transformation Initiative - Chemistry and Advanced Materials Industry (2017).

[14] ダイヤモンド社：フェイクニュース特集：ハーバードビジネスレビュー 2019 年 1 月号 (2019).

[15] 井上紫織，宇根正志：金融分野で活用される機械学習システムのセキュリティ分析，日本銀行金融研究所ディスカッション・ペーパー・シリーズ (2019)（オンライン），入手先 ⟨http://www.imes.boj.or.jp/research/papers/japanese/19-J-01.pdf⟩.

[16] 磯村　哲，山下博史：化学産業における分子デザイン，人工知能学会誌，Vol.34, No.3, pp.358–363 (2019).

[17] 科学技術振興機構（編）：研究開発の俯瞰報告書システム・情報科学技術分野（2019 年），chapter「機械学習における公平性・解釈性・透明性（FAT/ML）」，pp.109–117 (2019).

[18] 丸山　宏：高次元科学への誘い (2019).

[19] 原　聡：私のブックマーク：機械学習における解釈性 (2018)（オンライン），入手先 ⟨https://www.ai-gakkai.or.jp/my-bookmark_vol33-no3/⟩.

[20] 黒川通彦，平井智晴，櫻井康彰：データドリブン経営の本質，ハーバードビジネスレビュー 2019 年 6 月号，pp.21–35 (2019).

[21] 神嶌敏弘：公平考慮型データマイニング技術の発展，第 31 回人工知能全国大会，Vol.1E1-OS-241-1 (2017).

[22] 人間中心の AI 社会原則検討会議：人間中心の AI 社会原則 (2019).

浦本 直彦　（正会員）

1965 年生．1990 年九州大学大学院総合理工学研究科修了．同年日本 IBM 入社，東京基礎研究所に配属，自然言語処理，情報統合，Web 技術，情報セキュリティ等の研究開発に従事．2017 年三菱ケミカルホールディングスに入社．2018 年より人工知能学会会長および九州大学客員教授を兼務．人工知能学会，ACM 各会員．