

## AIに必要とされる透明性と説明可能性【前編】 G20が議論し始めた“人間中心のAI”

FUJITSU JOURNAL / 2019年8月28日



6月末に開催された20カ国・地域首脳会議（G20大阪サミット）には、そうそうたるメンバーが世界中から大阪に集結しました。二国間での首脳会談も積極的に行われ、日本の安倍晋三首相はホスト国の利点を生かし、米国や中国、ロシアなど多くの国と二国間会談を行いました。ただ最大の話題は米国のトランプ大統領と中国の習近平国家主席による米中会談だったかもしれません。貿易戦争真っ只中の両国の動向は、日本の多くの企業に影響するため、その成り行きに注目した方も多かったのではないのでしょうか。

もちろん、G20大阪サミットの首脳宣言にも注目すべきものが様々にありました。貿易面では、「自由、公平、無差別で透明性があり予測可能な安定した貿易及び投資環境を実現し、我々の市場を開放的に保つよう努力する」との文言が盛り込まれ、世界貿易機関（WTO）の改革に取り込

むことも明記されました。また、環境・エネルギー面では、プラスチックごみの流出による海洋汚染を2050年までにゼロにする「大阪ブルー・オーシャン・ビジョン」が盛り込まれました。

一方、デジタル面では、「データの潜在力を最大限活用するため、国際的な政策討議を促進することを目指す」と明記されました。また「人工知能（AI）の責任ある開発及び活用は、SDGsを推進し、持続可能で包摂的な社会を実現するための原動力となり得る」としました。AIを正しく活用することが、持続可能な社会を作るためには重要であるということ、各国の首脳が認識し合ったのです。

## G20大阪サミットで歓迎された「AI原則」とは？

実はG20大阪サミットに先だって2019年6月8～9日に茨城県つくば市で開催されたG20貿易・デジタル経済大臣会合においての、最大の話題はAIでした。この会合では、G20の貿易大臣とデジタル経済大臣が初めて一堂に会しました。そして、日本主導で閣僚声明をまとめましたが、その中に人間中心のAIを実現する環境づくりに取り組むとした「AI原則」を盛り込んだのです。

AI原則では、あくまでAIは人間の道具であるという前提で、無秩序なAIの開発や利用を防ぐことを推奨しています。そして注目すべきは、AIの開発者や運用者に「透明性」および「説明可能性」を求めていることです。AIの仕組みを正しく理解している人はほとんどいないでしょう。そのAIの出した判断が、仮に自分が望んでいないものだとする、どうしても不満や不安が残ってしまいます。そこでAI原則では、AIが下した判断に対して根拠を示すことを要求しているのです。

AI原則は、G20大阪サミットの首脳宣言でも触れられています。AIが人々から信頼され、かつその潜在能力を十分に引き出すために、AI原則を歓迎するとされました。

AIが社会の様々なシーンで登場するようになりました。AIがますます進化するのは間違いありません。そしてAIが社会においてより重要な意思決定に関わるようになった時、その意思決定に不満や不安を残してはなりません。必要となるのはAIの判断が本当に信頼できるものであるという根拠を示すこと。AI原則の採択は、それが世界中で求められているという証です。

AIによる判断は、ブラックボックスの中で処理されてはならず、透明性がなくてはなりません。そこで今回から2回にわたって、AIが下した判断の根拠を説明可能とする技術について見ていきます。

## AIが「ブラックボックス」では人々に不安が残る

現代は、「第3次AIブーム」といわれています。その立役者は、間違いなく「ディープラーニング」という技術が確立したことです。1980年代に訪れた「第2次AIブーム」では専門家の知識を記

述する「エキスパートシステム」が、技術の中心でした。ある専門家の知識を詰め込んで問題解決を行うエキスパートシステムは、一定の成果を残したものの、専門家の莫大な知識を詰め込むということが大前提だったために、コストも時間もかかるのが大問題。そのブームは長くは続きませんでした。

一方、ディープラーニングは大量のデータがあれば、コンピューターが自分で判断するルールを見つけ出します。ある程度のデータがあれば、いちいち専門家の知識を詰め込むなどの面倒な作業が必要ありません。このため、第3次AIブームは未だにその発展の途上ですし、実用面での応用が急速に拡大しています。今やAIは、決して手に届かないフェーズではなく、自社にどう活用していくかを真剣に考えるフェーズに来たともいえるでしょう。

ところが、人間が知識を詰め込んでコンピューターに判断させていた従来と比較して、ディープラーニングでは非常に複雑な処理を行うために、なぜその判断が下されたのか、根拠が極めて分かりにくいのが実態です。たとえその専門家がディープラーニングで下した判断を分析しても、どのように導き出したのかわからないことが多く、すなわちディープラーニングはブラックボックスの中で決定が下されるという問題が発生しています。

## 「説明可能なAI (XAI)」の研究が世界でブームに



この問題に対して、AI原則でも解決を求めるような方向で一致しました。また総務省が2018年にまとめた「AI利活用原則案」の10原則の中にも、「透明性の原則（AIサービスプロバイダ及びビジネス利用者は、AIシステム又はAIサービスの入出力の検証可能性及び判断結果の説明可能性に留意する）」「アカウントビリティの原則（AIサービスプロバイダ及びビジネス利用者は、消費者的利用者及び間接利用者を含むステークホルダに対しアカウントビリティを果たすよう努める）」として盛り込まれています。

世界的にもAIが下した判断に対して説明を可能とする研究が盛んになっています。このような技術は「説明可能なAI（Explainable AI、XAI）」と呼ばれており、NeurIPS（Neural Information Processing Systems）やICML（International Conference on Machine Learning）といったAIでは著名な学会でも、XAIに関する発表の数は増えてきました。例えばIBMは、判断を導くうえでの偏り（バイアス）を検知して手順を可視化する研究を行い、クラウドサービスの提供を始めています。また米国の新興企業であるシムマシーズは、データ同士の意味が似通っている度合いを数値で示すアルゴリズムを活用する研究を行っています。

富士通も、XAIに積極的に向き合っています。富士通では主に3つの技術を使ってAIのブラックボックス問題の解消に取り組んでいます。グラフ構造のデータを学習して推定因子を特定する「Deep Tensor™（ディープテンソル）」、情報同士の関係性を示す「Knowledge Graph（ナレッジグラフ）」、そして判断の仕組みが分かる学習モデルを備えた「Wide Learning™（ワイドラーニング）」です。（注1）

（注1） Deep TensorとWide Learningは富士通株式会社の商標です。

動画：[【AI】全ての組み合わせを網羅し特徴を見つけるAI - Wide Learning™](#)

全ての組み合わせを網羅し特徴を見つけるAI - Wide Learning™

Deep Tensor™とWide Learning™は、AIの中核と言える機械学習技術である一方、Knowledge Graphはそれを支える技術であり、両者は組み合わせで効果を発揮します。またWide Learning™は、これまで説明してきたディープラーニングとはまったく異なるアプローチでAIとしての判断を行います。このWide Learning™の詳しい説明は後編に譲ることとし、今回はDeep Tensor™ + Knowledge Graphの組み合わせが、AIの判断を説明可能にすることについて説明します。

## ディープラーニングを拡張する「Deep Tensor™ + Knowledge Graph」

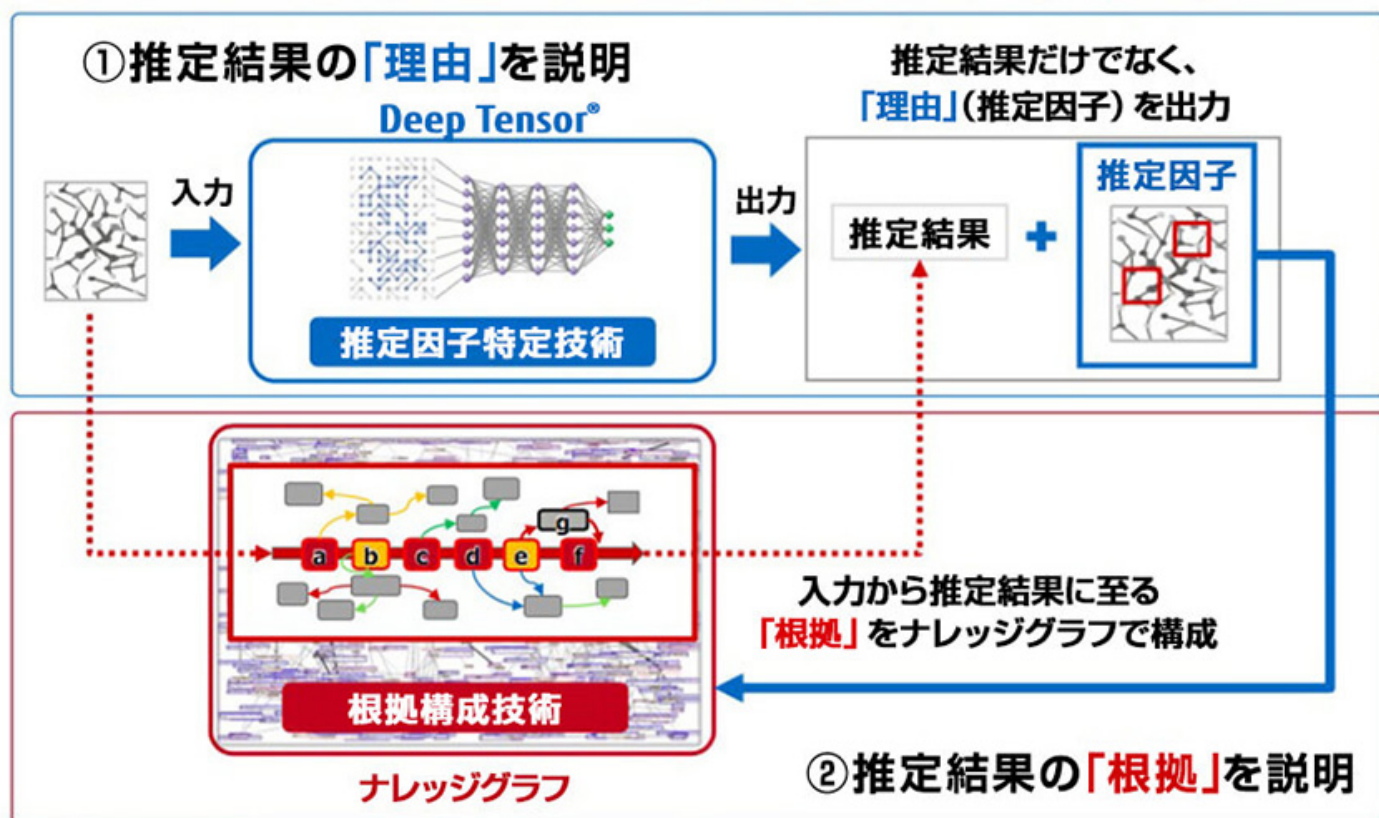
Deep Tensor™は、ディープラーニングを一步進化させた技術と言えます。導き出した判断と共に、その判断に大きく寄与した推定因子を特定できるのが最大の特徴です。具体的には、人やモノの「つながり」など関連を記述する場面で用いられるグラフ構造のデータ群を、「テンソル分解」と呼ばれる数学テクニックで変換すると同時にディープラーニングによる学習を行いま

す。さらには、ディープラーニングの出力結果を逆に探索することで、推定結果に影響した複数の因子を特定します。

一方のKnowledge Graphは、膨大なデータの意味や周辺知識の関連性に着目して、この関連性にもとづいて整理されたデータベースのことです。例えば、膨大なWeb情報や論文などを収集してナレッジグラフ化することで、知識体系をデータベース化できます。

Deep Tensor™が特定した推定結果に大きな影響を及ぼした因子と、Knowledge Graph上のデータを関連づけることで、それぞれの因子に関連性の高い情報を「根拠」として抽出できます。つまり、Deep Tensor™が特定した因子をナレッジグラフ上の具体的な知識と結びつけることで、ディープラーニングによる推定結果の理由や根拠を人が説明できるようになるというわけです。

## 2つの説明を行う全く新しいAI：結果の「理由」と「根拠」を説明



この技術の適用は、既に金融分野で始まっています。例えば、企業に対する投融資についての信用リスク判定。信用リスクの判定は、一般的には対象企業から提出された財務諸表や業績データを担当者が精査し実施します。しかし、投融資の対象が中小企業などの場合は、財務諸表が提出されていなかったり、提出されていても内容の信憑性に欠けたりといった課題がありました。

ここでDeep Tensor™ + Knowledge Graphの出番です。財務諸表に頼ることなく、金融機関と企業の間の実際の取引履歴のデータをAIで処理することで、信用リスクの判定を可能にしたのです。

いくらのお金をどこへ振り込んだなどの取引履歴データは、グラフ構造を持っているためにDeep Tensor™による解析が可能です。Deep Tensor™で信用リスクを判定するとともに、推定結果に貢献

した特徴的な因子を特定。それを金融機関にかかわる知識をもとに構築したKnowledge Graphと関連付けることで、理由・根拠の説明ができるようになります。信用リスクに問題なければ金融機関としては安心して投融資できますし、もし投融資を断らなくてはならない場合でも、断る理由が納得してもらえるようになります。

今後、自分たちの周りで判断をAIに委ねるケースがますます増えます。「え、なんでそうなの」と疑問に思うこともあるはずです。こんなとき、その判断に至った経緯を説明してくれば、納得する場合もあるかもしれません。[後編](#)では、富士通が開発したもう1つのXAI技術、Wide Learning™について詳しく見ていきます。

## 著者情報

### 木村知史

#### 日経BP総研 上席研究員

1990年慶應義塾大学理工学部機械工学科修士課程卒業、同年日経BPに入社。製造業のための総合情報誌「日経メカニカル」にて、主に先端の加工技術および生産管理システムの分野を取材・執筆。その後、CAD/CAMやSCMなど製造業におけるコンピューター技術の活用誌「日経デジタル・エンジニアリング」の創刊に参画。2008年1月、Webサイト「Tech-On！（現日経xTECH）」編集長。2012年7月、「日経ビジネス」編成長、2014年1月、同媒体のデジタル版「日経ビジネスDigital（現日経ビジネス電子版）」編集長。ビジネス関連のコンテンツを企画するとともにサイト運営やアプリ開発を先導。2019年1月、「日経ビジネス電子版」を企画・立ち上げ。2019年4月より現職。大手企業のコンテンツコンサルティングや「Beyond Health」等のメディアのコンテンツ企画を担当。