

AIアドバイザーリーボード勉強会（第3回） 議事抄録

日時：2021年12月7日（火）10時～12時

場所：豊洲センタービル10階 Inforium Lab B、Zoom

出席者（専門委員）：森川座長、石川委員、三部委員、奈良委員、成原委員

出席者（NTTデータ）：技術革新統括本部技術開発本部デジタル社会基盤技術センタ 市原主任、ヘルスケアAIセンタ 上子主任、総務部サステナビリティ担当 金田シニアスペシャリスト、品質保証部 中野課長、デジタル社会基盤技術センタ 鈴木シニアエキスパート（聴講者は省略）

プログラム

| 時刻 | アジェンダ | 時間 | 発表者 | |
|-------------|-----------------------------------|-------------------------|------------------------|-------------|
| 10:00～10:05 | 趣旨説明、本日のプログラム | 5分 | 技術開発本部 鈴木SX | |
| 10:05～10:10 | 座長より開会のご挨拶、当社出席者の自己紹介 | 5分 | 東京大学 森川先生 | |
| 10:10～11:00 | 講演 AIシステムの品質と倫理 ～開発者・技術者の立場から～ | 50分 | 国立情報学研究所 石川先生 | |
| 11:00～11:20 | NTTデータの取組み紹介 | 信頼できるAI提供に向けたAI品質管理の取組み | 10分 | 技術開発本部 市原主任 |
| | AI画像診断の実用化に向けた取組みと課題 | 10分 | 技術開発本部 上子主任 | |
| 11:20～11:55 | 質疑応答・意見交換 | 35分 | ファシリテーター： 東京大学 森川先生 | |
| 11:55～12:00 | クロージング・集合写真撮影 | 5分 | 全員 | |

議事概要

（1）趣旨説明・本日のプログラム

事務局を代表してNTTデータの鈴木シニアエキスパートから、AIアドバイザーリーボード勉強会の位置づけと開催趣旨の説明、および本日のプログラムの紹介があった。

（2）座長挨拶・出席者の自己紹介

座長の森川先生から本日の勉強会開催に向けた挨拶が行われた。

続いて、NTTデータの出席者から簡単に自己紹介が行われた。

(3) 講演：AIシステムの品質と倫理 ～開発者・技術者の立場から～

専門委員である石川先生から「AIシステムの品質と倫理～開発者・技術者の立場から～」と題して、講演をいただいた。

まず、ソフトウェアやシステムの品質に関して、可用性や性能など様々な品質特性があり、そのシステムの重要性に応じて求められる指標を検討する必要がある。その検討の指針として、これらを具体化・定型化した汎用的なガイドラインが作られており、国内では「非機能要求グレード」、国際的には「SQuaRE」という標準があることを紹介された。

これに対して機械学習型のAIに関しては、2018年ごろから品質について検討が進められ、国内では自動運転など5つのドメインに応じて品質面からの確認項目をまとめたAIプロダクト品質保証コンソーシアムの「QA4AI」、および、標準に近い形で品質特性を定義してまとめた(国研)産業技術総合研究所の「AIQM」の2つがガイドラインとして策定されていることを紹介された。また、AIシステムの開発にあたり、契約と品質特性に関して整理したディープラーニング協会が作成した契約ハンドブックがあり、これには品質特性に関して契約上、約束できないことがあることもまとめられていることを紹介された。こうした検討状況を踏まえ、AIには不確かさ・不完全さがあることを原則に、その適用領域に応じて品質面での個別の具体化が必要なことを述べられた。

次にAI独自の最近のトピックとして、説明可能性と公平性について紹介があった。

説明可能性については、ディープラーニングはどういった計算をしているかは明確なもの、構造が複雑すぎて人による解釈ができなくなっている。そのため、医療や融資判断のような人による最終判断を必要とする場面において、なぜAIがそのような出力をしたのかを人が理解できる形で示すことが求められるようになり説明可能なAI(XAI: explainable AI)の研究開発が進められてきたことを紹介された。しかし、技術の開発が進むにつれて、例えば画像認識では学習のさせ方次第では、人が見たら何が映っているのか分からない画像でも特定の断片があるだけで犬であると回答することがあることや、説明を受ける人も実態を正確に表現していなくてもそれらしい説明になっていけば納得してしまうという懸念があることを紹介された。説明可能性については技術的には確実に進化しているものの、技術の現状を踏まえ、安易に説明を求めたり付与したりするのではなく、場面に応じて個別に判断する必要があることを述べられた。

公平性については、AIによる採用の判断や顔画像認識において、性別や人種による差別の問題が顕在化したことから注目を集めるようになり、学習データの偏りが主な原因であ

ることから技術面からの対策は見えてきている。しかしながら、機械学習の仕組みとしては少数例外的なデータを入れて学習すれば予測性能が低下することがあることから、本当に公平性を担保しようとするれば予測性能とのトレードオフへの対応が必要となる場合もある。そのため、実際には場面に応じた公平の考え方やAIの出力を受けた意思決定の在り方について、プロダクトオーナーやステークホルダーも巻き込んで考える必要があることを述べられた。

最後にまとめとしてAIガバナンスの確立に関して、従来のソフトウェア開発ではオーナーの柔軟な要求を具体化してソフトウェアに実装するところまでを技術者を中心に進められてきたが、AIを活用する場合には、これまでに述べたようなAI独自の課題への対応も必要なことから、法律や社会学の専門家の方々や対象分野の専門家など多様なステークホルダーと技術者が協働して「具体化」を行っていく意思決定が必要であることを述べられた。

(4) 講演：NTTデータの取組み紹介 信頼できるAI提供に向けたAI品質管理の取組み

NTTデータの取組み紹介の1件目として、技術開発本部の市原主任から信頼できるAI提供に向けたAI品質管理の取組みの紹介があった。AI品質アセスメントとAI品質コンサルティングの2つがあり、AI品質アセスメントはこれまでのAIシステム開発の知見に基づいて、そのAIプロジェクトにはどういったリスクがあるのかを評価・分析するものであり、発生するリスクを未然に防ぎながらプロジェクトを進めることが期待できる。後者のAI品質コンサルティングは研究開発中の取組みで、AI品質アセスメントサービスで可視化したリスクについて、具体的な対策の立案やツール導入等の技術支援を行い、品質向上に貢献することを目指している。

(5) 講演：NTTデータの取組み紹介 AI画像診断の実用化に向けた取組みと課題

NTTデータの取組み紹介の2件目として、技術開発本部の上子主任からAI画像診断の実用化に向けた取組みと課題の紹介があった。前半はAI画像診断技術について、CTやMRIなどの撮影機器の進歩により1回の撮影で取得できる画像が数百枚単位に増えているものの、撮影された医用画像を見て診断する放射線科医は不足したままであり、「読影」と呼ばれる画像を読む作業をAIによりサポートすべく研究開発が進められていることが紹介された。後半は実用化に向けた課題について紹介があり、最終的な判断の責任を医師が負う必要が

あることからAIで実現できる範囲が限られていること、医療機器の商用化には薬事承認を得る必要があるが承認までにかかる期間が長くAIの技術進歩スピードとの乖離が大きいこと、データセットの構築には専門家の知見が欠かせないがコスト面で課題があること、の3点が指摘された。

(6) 質疑応答・意見交換

座長の森川先生のファシリテートにより、講演に対する質疑応答・意見交換が行われた。主な論点は以下のとおり。

【鈴木シニアエキスパート】

石川先生の御講演に関して1つお願いしたい。今日は最新の技術動向を御紹介いただいたが、最後のまとめスライドを見ると、まだ多少の課題があるようにお見受けした。最先端でこの分野を見ていらっしゃる見地から、今後の進歩の見通しをどのように感じられているかお伺いしたい。

【石川委員】

技術の方向性は結構、確立されたように感じている。研究者としては品質を説明するために必要な技術などが整ってくると良いと思う。個人的な課題感としては、技術の発展よりも、使いこなすや技術が専門ではない人との対話といった方に課題が移ってくるのではないか。

【森川座長】

ありがとうございます。何か予想もつかないようなブレークスルーが起これば別だが、そうでなければ既に先は見えているということか。

【石川委員】

新しい話が突然ブレークする可能性はゼロではないが、それほど高くない気がしている。XAIや公平性の成果が急激に増えて、大手のITベンダからもこれらのツールを揃えたというアナウンスが出ているので、これから流行っていく感じはする。しかし機械学習型のAIでは、大きなパラダイムシフトやこれまでに全く出ていない概念がこれから出るかという

と難しいように思う。強化学習や実世界で行動させるロボットなどでは、品質の考え方から発展して自律性や制御可能性といったブレークスルーがあるかもしれない。しかし機械学習型は、コンセプトは出揃ったが現実にはやるとなると大変なものも多くあり、実際には何をどこまで本当にやるのかといった方向になっているように感じている。

【森川座長】

ありがとうございます。他はいかがですか。

【市原主任】

石川先生の御講演にあった品質に関する意思決定についてお伺いしたい。

実際に我々がお客様と一緒にシステム開発をする際に、目標とする品質についてプロジェクトの始めに合意を取る必要がある。しかし、AIの開発ではリスクがあり、最終的に目標とした品質に達するかどうか確率的になってしまい達成できないケースも多々あるように思う。そのような場合に、具体的な目標値やメトリクスをどの程度しっかりと握っておくべきなのか、逆にどういったプロセスで考えるべきなのかアドバイスをいただきたい。

【石川委員】

確かに最初から約束するのは難しいが議論は必要。AIの正解率だけでは決められないので、考え方としてはAIを適用する組織、例えば工場であれば歩留まりなどの指標や導入コスト、削減できるコストとAIの正解率を結びつけるような形になるのではないかと。例えば製品を検査するAIの正解率が90%であれば、不良品にも関わらず正常と判定されて誤って出荷される製品がどのくらい発生して、それを正常な製品に交換する追加コストと企業として失う信用リスクがどのくらいか、さらには導入のコストと工場全体で抑えられるコストがどのくらいかといったことが最初から見えるのが理想なのではないかと。

【中野課長】

石川先生の御講演で、データ、モデル、システムに加えて、試行錯誤して改善を継続・反復できるアジリティーを持ったプロセスが重要というお話があった。これは学習データをもう一度、品質や公平性といった観点で見直して、その結果でまた試行錯誤するといったアジリティーが重要ということか。

【石川委員】

ご認識のとおりである。データの良し悪しはデータだけ見ても分からないところもあるので、そこから規則性を学習できるかを試した上で、さらに公平なのかを見ると、こういったデータが実は少なかったと後から気づくことは普通に起きる。最初から全てを見据えるのは難しいので、そこを何度も実験する速さであるとか、そのときに常に様々な専門家がいたりかといったことがアジリティーということで、QA4AIのガイドラインに挙がっている。

【森川座長】

ありがとうございます。オンラインで入られている奈良先生、成原先生、何かございますか。

【成原委員】

2点質問したい。先ほどの石川先生の御報告の中で、説明可能性について、機械学習・特に深層学習については説明が困難と言われているところ、XAIの開発が試みられているが、決定木を使うなど、機械学習・深層学習とは異なった仕組みを使って近似的に説明をしているといったお話があった。その際に実際のメカニズムと説明が違うことによる弊害は生じているのか、また、それに対してどのような解決策があるのかをお伺いしたい。もう1点、関係して、先ほどのAI画像診断のご報告について、AI画像診断が今後重要な役割を果たすことが期待されると思うが、その際に説明可能性や透明性といった課題がどのように問題になっているのか。医師が最終的に判断して説明するので、それほどAI自体の説明可能性や透明性は問題になっていないのかどうかをお伺いしたい。

【石川委員】

説明が違う弊害はあるはずで、例えば説明を見て適切な判断だと思ってAIを信じたら、説明が間違っていてAIの結果もたまたま間違っていたというケースなどが弊害になる。しかし、本当に難しいのは、それが起きていることを把握するには、調べようと思わないと調べられないところにある。実際の運用の中で全ての記録を残して、ある1日の結果を全て見直すようなことをやらないと誰も気づけない。そのため、説明が違った弊害が起きた

という実例の報告例はないのではないかと。人工的に作ったデータを使って変な説明を出させるといった研究はあるが、実際のシステムで説明が悪かったことでこのような問題が起きたといった事例は私の知る限りはない。起きていないのもあるかもしれないが、起きたことを把握するには、それを実験だと思って色々と調べないと分からないのが大きい。

【上子主任】

AI画像診断でも透明性は問題になってくるが、AIで「この辺りが怪しい」と示すことさえできれば、AIの判断基準がどうであろうと放射線科医の先生方はそれを見て独自に判断ができる。放射線科医ではなく、他科の専門医や読影に不慣れな先生方にはより踏み込んだ説明が必要になってくる可能性はあるが、放射線科の先生方からは「怪しい場所を示してくれれば、こちらで判断できるから大丈夫」というお答えを頂くことが多いので、そこまでプロダクトにまだ求められていないという感覚である。

【成原委員】

AIの透明性・説明可能性がどれくらいクリティカルになるかは、それを受け取る側の人間にどれくらいプロフェッショナルな判断が期待できるかということも関わってくるように思う。ありがとうございます。

【森川座長】 ありがとうございます。では奈良先生、お願いします。

【奈良委員】

石川先生のご講演にあった、具体化がとても重要であり、抽象と具体の行き来には複数の専門家を含めた多様なステークホルダーが関わるのが重要というお話と、説明可能性についての説明をめぐるアメリカの標準化団体NISTによる4つの原則も、私の専門であるリスクコミュニケーションの原則と非常に近いものがあり、どちらも納得いくものであった。質問としては、AIの出力に対する説明に関するもので、今後、AIがさらに普及や定着に至るには、2段階のユーザーに対するAIに対する理解やリテラシーが必要だと思う。すなわち1段階目のユーザーは、医師や不動産屋、お金を貸す人、人事担当者などAIを使う人で、2段階目のユーザーは、患者さんや家を借りる人、ローンを組みたい人、その会社に採用されたいと思っているなどAIによって診断や判断をされる人になる。こういった

場合に「なぜ落ちたのか」あるいは「このAIはどういった仕組みなのか」を2段階目のユーザーから1段階目のユーザーが問われた場合にどのように説明したら良いかといった質問を開発側としてされたことはあるかどうか。また、そういった場合に、1段階目の人はどのように対応すれば良いのか、その辺りのお考えがあればお伺いしたい。

【石川委員】

AIになったから説明しなくてはいけないというよりは、人対人の場合でも、例えば賃貸や採用で不可となった理由を説明しないといけないが、本当の理由を全て説明すると生々し過ぎるので表現を緩めたり、理由をわざわざ言う必然性もないので言わない、言葉で気持ちを伝えて納得してもらうというように、ある意味、曖昧な形でやってきたところもあるように思う。それは、おそらくAIになっても変わらないかみしれず、逆に人間がやっていたときに、そこが本当に出来ていたのであれば、その指針を引き継いだAIを整えることができるのではないかと思う。AIと人だったらAIを信用しないという意識は必ずあるので、そのためにAIに説明可能性を求めるとしたら、人だけで動いているときにも納得はできていたのか、透明性はあったのかという話のような気もしている。

【市原主任】

具体的なビジネスレベルにまでは至ってないが、実際、金融や人事分野での利活用を考えてヒアリングしたことはあり、AIを使った採用では、落ちたときにその理由を聞かれるのは確実に想定できるので、それは考えておくべきとのことであった。

一方で、現在のAI活用はAIが完全に決めるのではなく、あくまで人事や金融の方にインサイトを与えて最終的に人が判断するプロセスがあるので、現段階では説明責任を果たせると思っている。しかし、今後AIの性能向上と社会への浸透が進んだ際に、AIが答えたときに「AIに何で私のことが分かるのか」といった反応をされるのは容易に考えられるので、そういったときにどういった観点を提示すべきなのか、具体的に何を返すべきか、改めてここに参画されている皆様と御相談させていただければと思っている。

【上子主任】

患者から医師に「AIはこれで正しいのか」と問われたらどうしようということだが、現状は、最終的には医師が自身の責任で診断をしてくださいということで、AIに責任転嫁を

してはいけないことになっている。しかし、今後さらにAIが発展して、例えば単独でAIが動く未来が到来したとしたらどうなるか。医師相手であれば、何かあった時には、判断ミスや見逃しがあったということで訴訟することになるが、AI相手にそれができるのか、AIの開発元が訴えられる側になるのか、今後どうなっていくのか気になるところである。

【森川座長】

ありがとうございました。石川先生のお話を伺っていて、しっかりとしたものは簡単には出来上がらないのでアジャイル的にやっていく必要があるように思った。

そしてNTTデータの皆さんへのお願いとしては2つある。1つはこういったことを議論する学会を主導して作ってはどうか。例えば製造業では、品質工学会のように日本科学技術連盟がやっていて、大手自動車メーカーを筆頭に品質の関係者が全国から集まった巨大な組織があり長い歴史もある。AIガバナンスは大切なことなので、何か同じようなものが必要であれば、NTTデータがものづくりでいうトヨタさんのようにやってもいいのではないか。ぜひ検討いただきたい。

もう1つはAI画像診断のところで、制度や法規の課題の解決にはNTTグループがロビー活動をやっていただくというのも重要かと思うので期待している。ありがとうございました。

【鈴木シニアエキスパート】

本日はどうもありがとうございました。次回は、3月上旬頃の開催を予定しております。日程が決まりましたら御連絡を差し上げます。

(7) クロージング・写真撮影

最後に会場の出席者全員で写真撮影が行われた。

— 了 —