

インタラクティブなマッチングAIで 実現するトランスペアレントな人の 意思決定支援

中尾 悠里 大堀 耕太郎 穴井 宏和

あらまし

近年広がっているAI（人工知能）技術には出力の理由を人間に理解できるように説明することが難しいというデメリットがある。この問題に対処するために、説明可能なAI（XAI）の研究が進められている。しかし、人の意思決定を支援するためには、XAIに対してインタラクティブ性を付与することが必要である。このインタラクティブ性の実現のために、富士通研究所は福岡県糸島市役所様の協力のもとで移住地マッチングAIを開発した。これは、都市部のユーザーが自分に適した移住候補地を選択するための意思決定支援を行う推薦システムである。このシステムで、ユーザーは自身の好みを入力すると、移住候補地が提示されるだけでなく、その候補地について学習できる。

本稿では、人との協働を実現するインタラクティブなAIについて述べる。更に、糸島市への移住希望者に対する被験者実験の結果を紹介し、今後の課題と方針について述べる。

1. まえがき

AI（人工知能）が第三次ブームを迎えて久しい。現在は深層学習を筆頭に収集したデータを分析し、予測・意思決定などに役立てるデータ駆動のAIが主流である。これらのAIには、人間が明に言及できない特徴量を抽出できるなどのメリットと、出力や特徴量が人間の直観に反するなどのデメリットを併せ持つ。ビジネスや日常生活の文脈では、AIが出した答えの当否だけでなく、人間の感覚に近いかたちで理由が説明される必要があるが、それが難しいという問題があった。

この問題を解決するために、説明可能なAI（XAI: Explainable AI）の技術が開発されてきた [1]。XAIは、AIの出した結果の理由を人間にも理解可能なかたちで出力することを目的としている。例えば、画像の中のどの部分が判断材料として有効だったのかを明らかにすることが挙げられる [2]。

しかし、ビジネスや日常生活における意思決定の支援のためには、単にXAIを開発することを目指すだけでは不十分である。ユーザー自身がAIの出力に対して批判的なフィードバックを与えて結果を変え、その相互作用の中でユーザー側も熟考を重ねた上で、実際の行動に移るアプローチが必要である。

これによって、AIは初めて人間と協働することができる。図-1は、このアプローチと従来のアプローチの違いを示したものである。従来は、人をいかにして説得できるかに焦点が当てられていた。一方、我々のアプローチの特徴は、ユーザーが批判的なフィードバックを与えることで、AIからの出力を基に良く考えた上で意思決定を行える状況を担保することを狙う点である。

この状況を実現するために、AI技術自体もユーザーが理解できるように構築される必要がある。AIの中身が見えるようにインターフェイスや技術そのものが構築されていることを、トランスペアレント（透明）であるという。インタラクティブなXAIを開発し、人間が試行錯誤のもとAIの仕組みを理解しつつ意思決定できるようにすることで、トランスペアレントで信頼できるAIの実現につながる。

本稿では、AIから受けた説明がユーザーにどのような影響をもたらし、ユーザーがどのように反応するか調べるために、富士通研究所と福岡県糸島市が共同で開発・実証実験を行った移住地マッチングAIの事例を紹介する。まず従来技術とその問題点を述べ、移住地マッチングAIを例として、上述のアプローチがどのように実際のサービスとして実装されるかを示す。その後、システムの評価実験の内容に関して述べ、今後の方針に関して記述する。

従来の人工知能のアプローチ



富士通研究所が目指すアプローチ

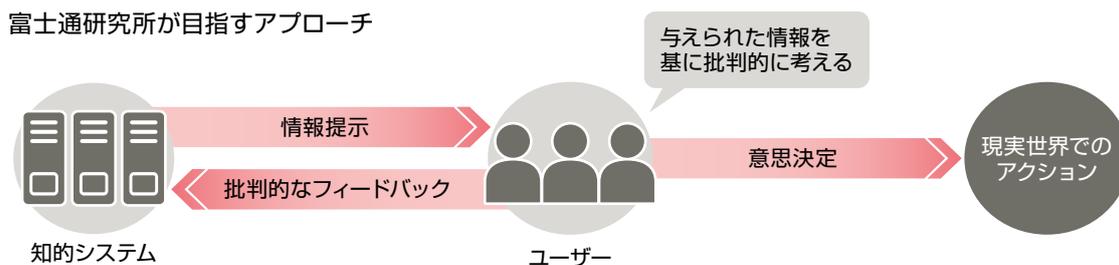


図-1 従来のアプローチと本アプローチとの違い

2. 従来技術とその問題点

知的システムとのインタラクションを通じて、人の意思決定を支援するためのシステムが存在する。学術的には、2000年代後半に、自動車やデジタルカメラといった商品の推薦システムにユーザーからのフィードバック機能を追加した、クリティキング (Critiquing) と呼ばれる技術が開発された [3]。また、ビジネスの文脈ではジョブマッチングや不動産マッチングなどのサービスにおいても、推薦システムに対して自身が設定した条件でフィルタリングするようなサービスが一般に広がっている [4]。これらのシステムは、ユーザーからのフィードバックを受け付けるが、技術自体をトランスペアレントにすることは重視していない。意思決定の内容が重要であればあるほど、なぜAIがその提案をしているかをユーザーが深く理解する必要がある。そのためには、技術自体のトランスペアレンシーを上げることが不可欠である。今回紹介するシステムでは、AI内部の仕組みもできる限りユーザーにとってトランスペアレントな状態でインタラクティブなシステムを開発した。

3. 人の意思決定を支援する 移住地マッチングAI

本章では、移住地マッチングAIというインタラクティブな推薦システムを紹介する。これは、日本国内の都市部に住み、地方への移住を希望する人が、自分がどのような地域に住むべきかを探すためのサービスを提供するシステムであり、福岡県糸島市役所、富士通研究所の共同研究によって開発された。都市部の住民は、地方自治体の事前知識を持たないため、実際の状況を学習しつつ自分が移住したい地域を探してもらう必要がある。このサービスでは、移住希望者に対して、糸島市内の行政区 (町内会) の中から住むのに適した地域を推薦する。富士通研究所は、はじめに既に都市部から糸島市へ移住した人、および今後移住を考えている人に対して複数回のインタビューを実施した。これによって、都市部の住民が移住地域を決定する際に考慮する項目

を洗い出した。そして、地域データを用いてそれらの項目を町内会ごとに定量化し、交通、買い物、病院などの地域指標を作成した。同時に、インタビューを基に、車の運転の有無や子どもの年齢層などの要因を用いて、移住希望者を数十のカテゴリーに分類した。更に、各々のカテゴリーで各々の地域指標がどの程度重視されるかを、定量的なアンケートを用いて割り出した。以上から、図-2に示すGUIを備えるシステムを開発した。

このGUI上では、移住者のカテゴリーが上部左側に表示され、そのカテゴリーの人が重視する項目 (地域指標) が上部右側に表示される。実際にどの地域指標をどの程度重視しているかは、図-2のStep1の棒グラフで示される。ユーザーは棒グラフの上下の矢印をクリックすることで、自分の意思と異なる地域指標の値を変更できるため、希望するバランスに調整できる仕様になっている。図中には表示されていないが、その下には移住先地域の推薦ランキングが、その地域の地域指標の値とともに表示されている。ユーザーは、それらの情報を見ながら自分の好みを表す棒グラフを調整していくことができる。

このシステム内部には、ユーザーとのインタラクションを学習する仕組みが実装されており、学習によってユーザーのカテゴリーを自動的に修正していく。これによって、ユーザーはシステムから提案された地域の情報を学習し、その一方で、より多くのユーザーが使うほどAIはユーザーの特性を学んでいく。また、本システムではシステムのトランスペアレンシーを確保するために、人間の意思決定の初期段階をモデル化したMulti-attribute utility theory [5] に則り、線形のモデルを利用している。つまり、各町内会のスコアは、地域ごとに付与された各地域指標の値と、棒グラフで表示されるその指標に対するユーザーの重視度の積を単純に足し合わせたものである。

4. システムの評価

本システムの効果を被験者実験で調査した。移住イベントに参加中の被験者16名、およびインターネットを通じて募集した、地方への移住希望はあるが糸島市のことは知らない被験者83名の合計99名

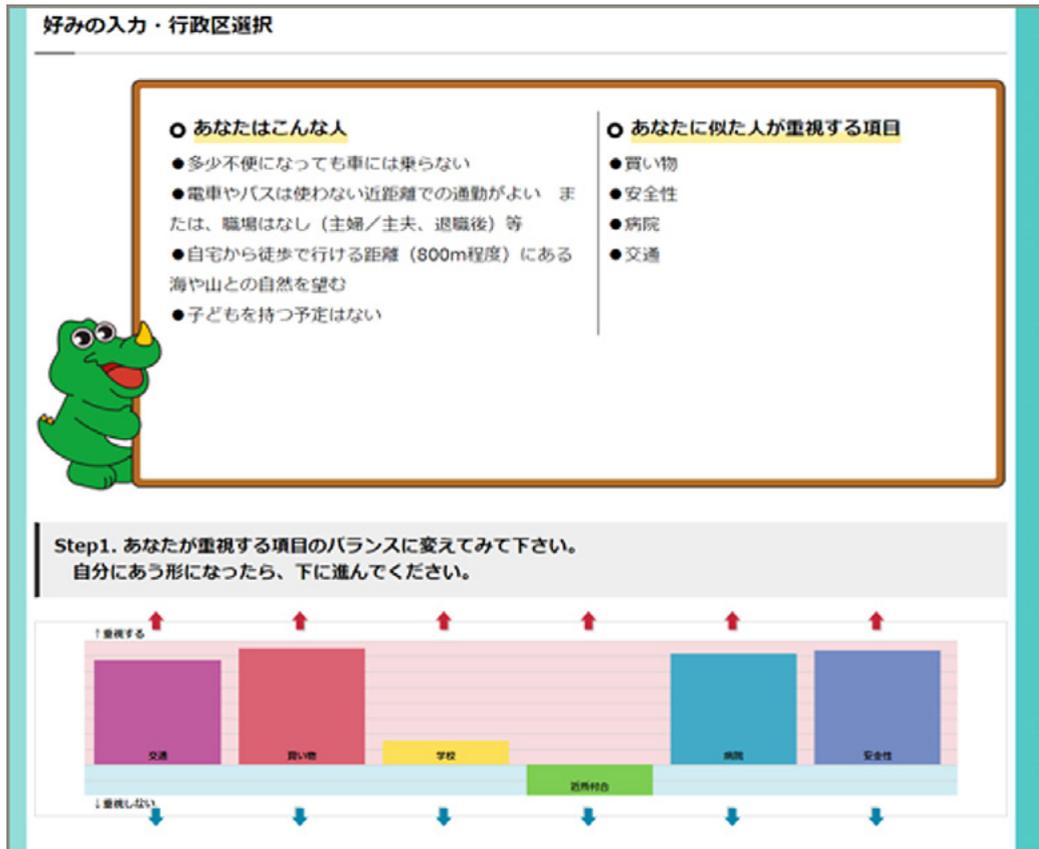
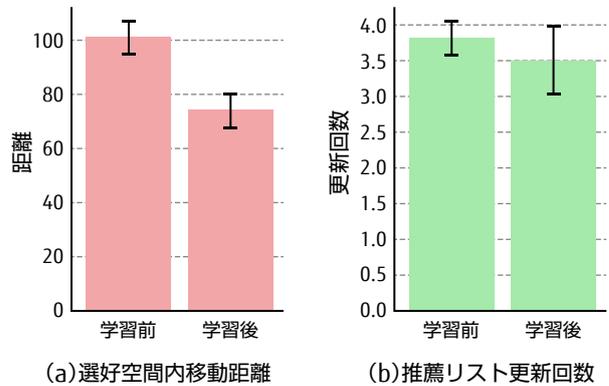


図-2 糸島市移住マッチングAIシステムのGUI

に対して被験者実験を行った。インターネット募集の被験者のうち、55名はユーザーカテゴリーなどの学習を行っていない状態のシステムで実験し、残りの28名は前半の55名分のデータを学習させた状態のシステムで実験した。評価内容として、学習前後で、実際にユーザーが移動した選好空間内の距離の変化、およびユーザーが推薦地域リストについて自分の好み（図-2中の棒グラフの長さで示されている、重視する項目に対する値）を変えることで更新した回数を調べた。選好空間とは、ユーザーが各地域指標に対して与える重要度をプロットした空間である。今回地域指標は六つ用意したため、ユーザーの好み、すなわち選好は6次元の空間内にプロットされている。

まず、移住イベントに参加中の被験者は自身の好みを変えたい傾向があることがわかった。ある参加者からは、システムからは「ここに住め」と命令されているように感じるとのコメントがあった。



注)エラーバーは標準誤差

図-3 選好空間内移動距離と推薦リスト更新回数

インターネットで募集した被験者にも、システムを利用した移住地推薦を行い、その後で糸島市役所の担当者が移住相談会を実施するというフローを体験してもらった。図-3は、定量的な結果を示したものである。図-3 (a) の選好空間における移動距離

の平均値を、学習前と学習後で比較したものである。学習前の平均値は101.1であるのに対し、学習後の平均値は74.2でp値 [6] は0.003であり、有意に減少している。しかし、図-3 (b) の推薦リストを更新した回数については、学習前の平均値は3.8なのに対し、学習後の平均値は3.5でp値は0.55であり、有意差がない。被験者の数が学習前は55名、学習後は28名と大きく違いはするが、選好空間内を移動する距離が短くなっても、被験者が推薦リストを更新する回数は少なくならないという傾向がある。つまり、被験者は仮に自分の好みの地域がすぐに表示されたとしても、平均して4回弱は自分の好みを変更しようとする事がわかる。これは、人はより自分に合った地域を探すために、一定数は必ず推薦リストを更新したいと思うという傾向がある可能性を示している。

また、同じ条件下で得られた被験者の発言やアンケートへの自由記述からは、AIでの移住相談の良い点として、「人に相談するよりも気軽に早い」「気付きにくい地域にも気付くことができそう」「公平・客観的な情報を与えてくれる」といった点が挙げられた。一方、悪い点として、「実際の生活の中での話を文字情報から感じ取るのには限界がある」といった点が挙げられた。被験者は、糸島市役所の移住担当者からの案内も受けたが、この案内に対しては「多方面から、生活に関して具体的な提案がある」「細かな質問にすぐに答えてくれる」などの良い点が挙げられ、悪い点は挙げられなかった。

本システムは実証実験後の2017年9月から2018年4月までの期間で、糸島市役所のウェブサイト上で糸島市移住AIマッチングシステムとして実運用が行われ、複数の移住が実現した。

5. 今後の課題と方向性

今後の課題の一つは、インタラクションを促すインターフェイスの開発である。ユーザーとAIが相互作用した上で意思決定できるシステムの構築を狙ったが、実際の被験者はAIが出した提案を「答え」のように捉える傾向が強く、相互作用には積極的ではなかった。そのため、インターフェイス自体

の工夫によってインタラクションを起こすような改善が必要である。

更に、線形モデルを使うことによってトランスペアレンシーを高める工夫を行ったが、ユーザーは線形のモデル自体を把握することにも難しさを感じたため [7]、モデルをどのように簡潔に説明するかも課題である。

第1章で紹介したインタラクティブなXAIを開発することで、人間の意思決定を適切にサポートできるAIが可能になるという設計指向は、今回紹介した移住地マッチングAIに限られるものではない。同様の方法論を用いて、観光・不動産・教育をはじめ様々な分野に応用できるマッチングAIを設計していくことも考えられる。

6. むすび

本稿では、ビジネスや日常生活での意思決定を支援するためのインタラクティブなXAIの必要性、およびそのプロトタイプとして福岡県糸島市と共同で開発した移住地マッチングのシステムとその実証実験結果について述べた。今後も富士通研究所は、人と技術の接点をうまく設計することで、人の意思決定を適切に支援するインタラクティブなXAIシステムの研究開発を進めることで、信頼できるAIを実現していく。

本稿に掲載されている会社名・製品名は、各社所有の商標もしくは登録商標を含みます。

参考文献・注記

- [1] A. Abdul et al. : Trends and Trajectories for Explainable, Accountable and Intelligible Systems: An HCI Research Agenda. CHI '18 Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Paper No.582 (April 21-26, 2018).
- [2] M. T. Ribeiro et al. : "Why Should I Trust You?" : Explaining the Predictions of Any Classifier. KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, p.1135-1144 (August 13-17, 2016).

- [3] L. Chen et al. : Critiquing-based recommenders: survey and emerging trends. *User Model User-Adap Inter*, Volume 22, Issue 1-2, p.125-150 (April 2012).
- [4] C. He et al. : Interactive recommender systems: A survey of the state of the art and future research challenges and opportunities. *Expert Systems with Applications*, Volume 56, p.9-27 (September 2016).
- [5] W. Edwards et al. : Decision Technology. *Annual Review of Psychology*, Volume 52, p.581-606 (February 2001).
- [6] 統計的仮説検定において、ある仮説の下で検定統計量がその値になる確率。本稿では、対応しない二群から得られた測定量の平均値が同じであるという仮説をたて、それがどれほど起こりえないかを示すために用いる。p値が小さいほど、二群の平均値同士が同じであることはあまり起こりえない。本稿では、p値が0.05を下回れば二群の平均値が有意に変わったと判断している。
- [7] Y. Nakao et al. : Generation of Hints to Overcome Difficulty in Operating Interactive Recommender Systems. *IntRS '19: Joint Workshop on Interfaces and Human Decision Making for Recommender Systems* (September 19, 2019).



穴井 宏和 (あない ひろかず)

株式会社富士通研究所
人工知能研究所
数式処理、最適化、人工知能の理論と
応用の研究・開発に従事。

この記事は、富士通の技術情報メディア「富士通テクニカルレビュー」に掲載されたものです。他の記事も是非ご覧ください。

富士通テクニカルレビュー

<https://www.fujitsu.com/jp/technicalreview/>



著者紹介



中尾 悠里 (なかお ゆうり)

株式会社富士通研究所
人工知能研究所
推薦システム、公平性配慮型機械学習、
アクセシビリティのユーザー観察の研究に従事。



大堀 耕太郎 (おおほり こうたろう)

株式会社富士通研究所
人工知能研究所
人工知能・数理技術による社会システム
デザインの研究に従事。