

# 持続可能なデータ収集を可能にさせるサービスプラットフォームの構築

## Development of service platform and eco system for sustainable data collection of AI

高岡 昂太<sup>\*1</sup>

Takaoka, Kota

本村 陽一<sup>\*1</sup>

Motomura, Yoichi

<sup>\*1</sup> 産業技術総合研究所 人工知能研究センター

National Institute of Advanced Industrial Science and Technology, Artificial Intelligence Research Center

Artificial intelligence (AI) has been making extraordinary progress. To keep developing the AI, reciprocal feedback-communication between AI and people in many use cases are important to collect sustainably data. Hence, this study aims to effectively build an AI platform from data collection to data analysis as well as the eco-systems in the field. The platform includes multiple applications for data collection, a cloud database to store data, probabilistic latent semantic analysis and Bayesian network as the AI by which people understand why predictions and recommendations are provided as a white-box. The platform can be easily customized and comfortably deployed for each use case depending on user needs. In the development trial phase, the platform has been deployed in several fields, such as vending machines in local communities.

### 1. はじめに

人工知能 (AI) は、機械学習とビッグデータ、計算機の高速化により適用可能範囲が広がってきた。しかしながら、新たなデータが持続的に収集されないと AI 自体の性能は向上しない。つまり、多くデータが収集され続けるような環境やエコシステムを醸成する必要がある。そのためには価値あるユースケースで早期に AI の実装および活用がなされ、さらに多人数、多頻度、広範囲に展開する必要がある。

そのため本発表の目的は、AI 学習を進化させるために、サービスを提供することで、データを持続的に収集するサービスプラットフォームの必要性と、エコシステムの例を紹介する。

#### 1.1 本発表の位置づけ

本発表の位置づけは、実社会問題へのチャレンジにおいて、AI を活用・応用するために必要不可欠なプラットフォームの提案、及びそれを活用するエコシステムに関する内容となる。そのため、本発表は、今後の人と AI が協調する一つの基盤モデルとなる(図 1)。

### 2. AI サービスプラットフォームの構築

AI を発展させるには、データプラットフォームは重要である。しかしながら、業務の中で不必要なデータ入力や、フィードバックもなく不明瞭なデータ収集の継続は難しい(プロメテウスの岩問題)。そのため、業務やサービスと連動し、データ収集からフィードバックまで一貫したサービスプラットフォームが必要不可欠となる。特にそのような AI を実社会に実装するには、透明性 (transparency)・説明責任 (accountability)・実社会へのポジティブなインパクトという 3 要因が重要と言われる(Cath et al. 17)。

#### 2.1 現場のニーズに合わせたデータ収集

サービスプラットフォームとエコシステムの構築には、まず現場のニーズ把握が重要である。例えば、実生活場面において、AI 活用が、検知・評価・省力化・能力開発・マッチングなど、どの分野で期待されるかの把握である。また、本研究では個々人

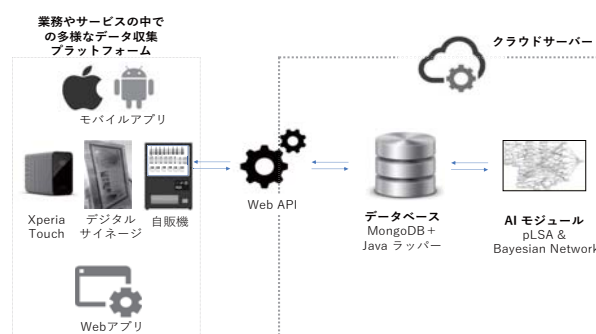


図 1: サービスプラットフォーム概要

の文脈や状況にあったフレームを理解するため、ユーザーモデル、現象モデル、確率モデリングという3つのモデル化を行った (Motomura, 2009)。以下にその内容を示す(図 2 参照)。

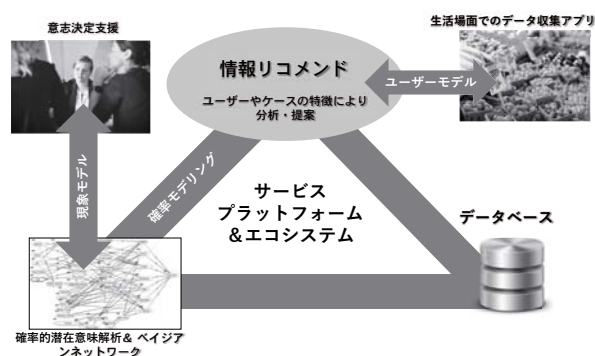


図 2. 場面ごとのモデリング

実生活場面において、ユーザーがどのような行動を取っているか、またその際にどのようなデータ収集がユーザーに負荷をかけずに可能になるかを検討した(ユーザーモデルの検討)。さらに、ユーザーがどのような場面で AI による意志決定支援をしてほしいのか、場面や文脈、実生活場面での制約といった、フレームに関する現象を検討した(ユーザーの意志決定支援を求める状況・現象モデルの検討)。そして、ユーザーモデル、現象モデルを事前情報として組み込み、リスク・コスト・ベネフィットを

連絡先: 高岡 昂太, 産業技術総合研究所 人工知能研究センター, 〒135-0064 東京都江東区青海 2-4-7, 03-3599-8920, kota.takaoka@aist.go.jp

最適化した確率モデリングを構築していった。確率モデリングについては、2.3 で述べる。

## 2.2 拡張性を想定したデータベース

実生活場面では、上記のユーザーモデル、現象モデル、確率モデリングを検討するために、解析目的上必要なデータ収集と実現可能なデータ収集、それぞれのニーズに応える柔軟なデータベースの拡張性が求められた。具体的には、取得される変数が増えることが想定される。データベースの一貫性を考慮することも重要であるが、サービスや現場実践と協働しながらデータ収集と解析を同時に行うサービスプラットフォームでは、現場のニーズに合わせて、柔軟に拡張できることに優先順位を置いた。その結果、サービスプラットフォームでは数値・テキスト・画像・動画などマルチモーダルなデータを収集するため、NoSQL 系データベースとして MongoDB を採用した。

## 2.3 ホワイトボックス化された AI 基盤

機械学習はデータから「予測」をすることが非常に強力である一方、実生活場面において、例えばビジネス場面で顧客への理由の説明や、公的場面での行政機関における説明責任など、「説明」が求められる場合は適さない場合がある (Executive Office of the President National Science and Technology Council Committee on Technology, 2016)。そのため、実生活場面における AI サービスプラットフォームと循環的に実生活にフィードバックしていくエコシステムでは、「説明」できることに重きを置いた。

また、サービスプラットフォームでは、データに基づき、ユーザーの行動をより良く変え、生活空間における質を高めることで、社会にポジティブな影響を与えることを目指している。そのため、ユーザーが主体的に AI を用いて判断するには、理解しやすく共感できる根拠の提示が不可欠となった。なぜならば、根拠が明確で受け入れられやすければ、ユーザーは近い将来の見通しを立てられ、自発的に行動を変える動機付けとなるからである。だからこそ、理由が明確で、理解しやすく、また実生活場面における様々な不確実性を考慮に入れられる確率モデリングを、AI サービスプラットフォームは採用した。具体的にはアルゴリズムは、基本的に確率的潜在意味解析(pLSA)(Hofmann, 1999)とベジアンネットワーク(Pearl, 1985)を採用している。

## 3. AI サービスプラットフォームとエコシステム例

同じユーザーでも、時間・場所・一緒にいる相手(フレーム)などにより、考え方や行動パターンは変化する。本発表において、単にシステムという呼び方ではなく、エコシステムとあえて呼んでいるのは、エコという血が通ったユーザーと AI との生態学的な関係性を重視しているからである。現在、ビジネス、地域活性化、社会問題にこの AI サービスプラットフォームの構築とエコシステムの醸成を同時に行っている (Takaoka, in press)。以下では、その一例として、開発した自動販売機でのサービスプラットフォームとエコシステムについて概要を説明する。

### 3.1 自販機サービスプラットフォームとエコシステム

自販機と言えども、飲料、食品、書籍類、日用品、電化製品など扱う商品は多岐にわたる。また設置場所により、ターゲットとなるユーザーも異なる。しかしながら、重要な点は、それぞれ地域コミュニティの中で、自販機は販売から災害時の商品配布、デジタルサイネージとしての情報啓発まで自律的な機能を発揮できることにある。例えば、本サービスプラットフォームを自販機に活用し、公共交通系 NFC カードと連動することで、ユーザーの特性に基づく商品を継続的にリコメンドできる。また、

設置場所などにおいても、購入商品が変わることが予想できる。例えば、“平日夕方頃”に“職場で一人で”お茶を買ってリフレッシュするユーザーは、“休日前”に“仕事上がりの夕方”に“同僚二人以上と一緒にいる”場面であれば、アルコールを購入する確率は高まる可能性がある。一方、同じユーザーでも、“平日 19 時頃”に“ジム”にいとスポーツドリンクを買う確率が上がる可能性もあるだろう。そのような個人に紐付くデータを、フレームとして事前情報を確率モデリングに組み込むことができれば、フレーム問題に対応しながら、リコメンドの理由をユーザーに示し、ユーザー特性に合わせて商品をリコメンドすることができる。つまり、サービス提供とデータ収集が循環的に可能となる。

このように、サービスプラットフォームとエコシステムは、人の主体性を根拠を持って支えるパーソナルガイドとして、サービスを提供し、その反応をデータとして持続的に集めることが可能になると期待される。

謝辞: 本調査の成果の一部は、国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) の委託業務の結果得られたものである。

## 参考文献

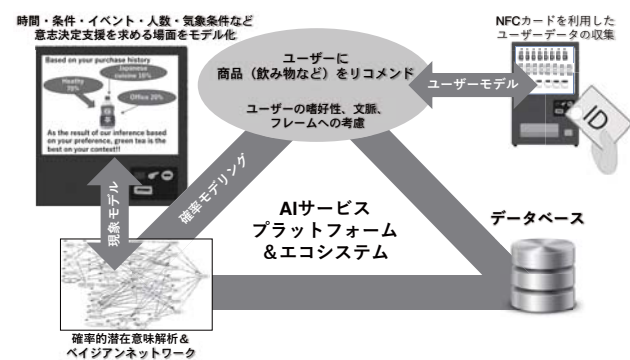


図 3. 自販機におけるサービスプラットフォーム

- [Cath 2017] Cath, Corinne et al. 2017. “Artificial Intelligence and the ‘Good Society’: the US, EU, and UK Approach.” *Science and Engineering Ethics*: 1–24.
- [Executive Office of the President National Science and Technology Council Committee on Technology 2016] Executive Office of the President National Science and Technology Council Committee on Technology: Preparing for the future of Artificial Intelligence, 10.1007/s00146-016-0685-0, 2016.
- [Hofmann, 1999] T. Hofmann: Probabilistic latent semantic indexing, *Proc. 22nd Annu. Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.*.
- [Motomura 2009] Y. Motomura: Predictive modeling of everyday behavior from large-scale data, Synthesiology English edition, National Institute of Advanced Industrial Science and Technology, 2009
- [Pearl, 1985] J. Pearl: Bayesian networks: A model of self-activated memory for evidential reasoning, *Proc. 7th Conf. Cogn. Sci. Soc.*.
- [Takaoka in press] Takaoka, K., Yamazaki, K., Sakurai, E., Yamashita, K., and Motomura, Y: Development of an Integrated AI Platform and an Ecosystem for Daily Life, Business and Social Problems, Applied Human Factors and Ergonomics (AHFE 2018), Springer, in press