



HOKKAIDO
UNIVERSITY

講義「人工知能」 第8回 AI関連ニュース

北海道大学大学院情報科学研究院
情報理工学部門 複合情報工学分野
調和系工学研究室 准教授 山下倫央

<http://harmo-lab.jp>

tomohisa@ist.hokudai.ac.jp

2024年5月2日(木)

❖ OECDが生成AI対策、開発者に偽・誤情報への対処求める...閣僚理事会で指針改定案採択へ

- <https://www.yomiuri.co.jp/economy/20240430-OYT1T50220/>

- 2024年5月1日 読売新聞オンライン

- 経済協力開発機構（OECD）が見直しを進めているAI（人工知能）に関する国際指針「AI原則」の改定案が判明した。

- 2019年の採択後、精巧な文章や動画を作り出す生成AIの利用が急拡大したことから、AI開発者らに対し、偽・誤情報への対処を求める項目を新たに盛り込む。

- 改定案は5月2、3日に開かれるOECD閣僚理事会で採択される見通しだ。

❖ OECDが生成AI対策、開発者に偽・誤情報への対処求める...閣僚理事会で指針改定案採択へ

- <https://www.yomiuri.co.jp/economy/20240430-OYT1T50220/>

- 2024年5月1日 読売新聞オンライン

- 改定案では、日本が昨年にG7（先進7か国）議長国として主導した「広島AIプロセス」の合意内容を反映して、AIの透明性や説明責任に関する項目も見直す方向だ。開発者らに対し、「AIの能力や限界に関する情報」や「AIの学習データや生成過程に関する情報」の開示を求める。

- OECDのAI原則は法的拘束力はないが、46か国が署名しており、国内施策の策定に影響力を持つ。

- ❖ ゴミ出しの問い合わせにA I が架空の部署案内、正答率は目標に届かず「市民向けには時期尚早」

- <https://www.yomiuri.co.jp/national/20240501-OYT1T50023/>

- 2024年5月1日 読売新聞オンライン

- 瀬戸内海に面し、製造業や農業が盛んな香川県 三豊みとよ市。環境衛生課課長補佐の岡崎英司さん（51）は昨年11月、役所のパソコン画面を見て、嘆息した。
- 視線の先には、実証実験中の対話型A I（人工知能）サービス「チャットGPT」が表示した文章。ゴミ出しに関する質問の答えに大きな間違いがあった。

- ❖ ゴミ出しの問い合わせにA Iが架空の部署案内、正答率は目標に届かず「市民向けには時期尚早」
 - <https://www.yomiuri.co.jp/national/20240501-OYT1T50023/>
 - 2024年5月1日 読売新聞オンライン
- 三豊市は昨年6月から、生成A Iに詳しい東京大学の研究室の協力を得て、ゴミ出しに関する問い合わせシステムの実証実験を始めた。
- 6万人余りが暮らす市には最近、技能実習生など外国語を話す住民も増えている。24時間いつでもゴミの出し方に関する市民の疑問に様々な言語で答えてくれる。そうすれば職員の負担が減る。A Iへの期待は大きかった。
- だが、「誤回答」が大きな壁となった。目指した正答率は「99%以上」。4000件を超える質問をA Iに学習させ、当初の63%から94%まで向上したが、それが限界だった。東大研究室の担当者は「今の技術では、正答率をさらに上げるのは難しい」と話す。

❖ 香川・三豊にAI研究拠点 地方の課題を飛躍の資源に

- <https://www.nikkei.com/article/DGXMZO41893610Y9A220C1LA0000/>
- 2019年3月1日 日本経済新聞
- 最新の人工知能（AI）研究を地方から発信する試みが動き始める。
- 香川県出身でAI研究に取り組む東京大学大学院の松尾豊・特任准教授が同県三豊市にサテライト研究室を設け、人手不足に悩む地元企業の作業効率化を模索する。
- 課題を多く抱える地方だからこそ研究対象が豊富にある。不利な条件を逆手に取り、飛躍の資源にする地域作りに期待がかかる。

❖ ChatGPTでの業務効率化を“断念”——正答率94%でも「ごみ出し案内」をAIに託せなかったワケ 三豊市と松尾研の半年間

- <https://www.itmedia.co.jp/news/articles/2312/15/news158.html>

- 2023年12月15日 ITmedia Inc

- 「ごみ出し案内」「市役所業務の中で、1番の懸念である個人情報漏えいの恐れがなく、かつ市民の利便性が高いものとして、ごみ出し案内は入り口が広がったため実験を開始した。しかし、廃棄物の形状が数多ある上に、地区やモノによって異なる収集日をかけ合わせると、かなり複雑な思考回路から答えを導き出さなければならず、結果的にごみ出し案内に使うには出口が狭いものであると痛感した」
(三豊市)

- また、誤った回答をしないようにAIが答えられない回答や、不安のある回答については「環境衛生課へお問合せください」という文言を表示していた。これにより“結局職員の負担は減らない”と分かったことも判断の理由になった。

❖ ChatGPTでの業務効率化を“断念”——正答率94%でも「ごみ出し案内」をAIに任せなかったワケ 三豊市と松尾研の半年間

- <https://www.itmedia.co.jp/news/articles/2312/15/news158.html>
- 2023年12月15日 ITmedia Inc
- 今回の実証実験は大きく2つのフェーズに分けられる。前半（6月1日～7月7日）では、約1週間のスパンで誤答を分析した。例えば「学習していない品目に対して誤答したケース（例：ファブリーズの捨て方など固有名詞があるもの）」「学習していたのに正しい回答ができなかったケース」「回答はできたが情報量が不十分だったケース」など正しく回答できなかった事例の改善を試みた。

❖ ChatGPTでの業務効率化を“断念”——正答率94%でも「ごみ出し案内」をAIに任せなかったワケ 三豊市と松尾研の半年間

- <https://www.itmedia.co.jp/news/articles/2312/15/news158.html>
- 2023年12月15日 ITmedia Inc

- この期間中のAIの正答率は62.5%という結果に。誤答の分析を試みても、正答率は大きく上がらなかったため、一度実証実験は中止することとなった。そこで改善点として、以下の4点を変更した。
 - 大規模言語モデルをGPT-3.5からGPT-4に変更
 - 1問1答形式から対話形式に変更（不足情報があるとAIが判断にした場合、ユーザーに対して質問する）
 - AIによる推測情報の表示（AIの推測が間違っていた場合、ユーザーからの指摘を促し、AIが回答の修正を行うことを期待）
 - 返答のリアルタイム表示（GPT-4に変えた際、やや応答が遅くなったため、回答を全て生成してからの表示ではなく、リアルタイム表示にして、待ち時間の体感が短くなることを期待）
- これらの改善を終えた10月23日から実証実験を再開。11月30日まで検証したところ、正答率は94.1%まで向上した。

- ❖ ChatGPTでの業務効率化を“断念”——正答率94%でも「ごみ出し案内」をAIに任せなかったワケ 三豊市と松尾研の半年間

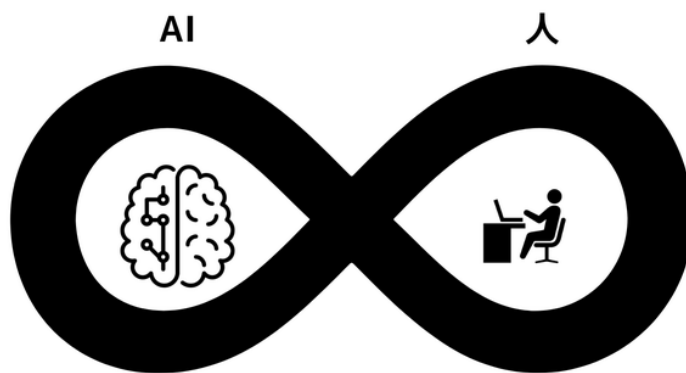
- <https://www.itmedia.co.jp/news/articles/2312/15/news158.html>

- 2023年12月15日 ITmedia Inc

- 「AIには少なくとも職員と同等のレベルを求め、それに達しない限り対市民向けとしては導入できないと考えた。また、AIがどのように回答したかを結局のところ職員が確認する作業が伴い、正答率が低ければそれだけ確認する頻度も上げなければならない。100%は無理としても99%は譲れない条件だった」（三豊市）

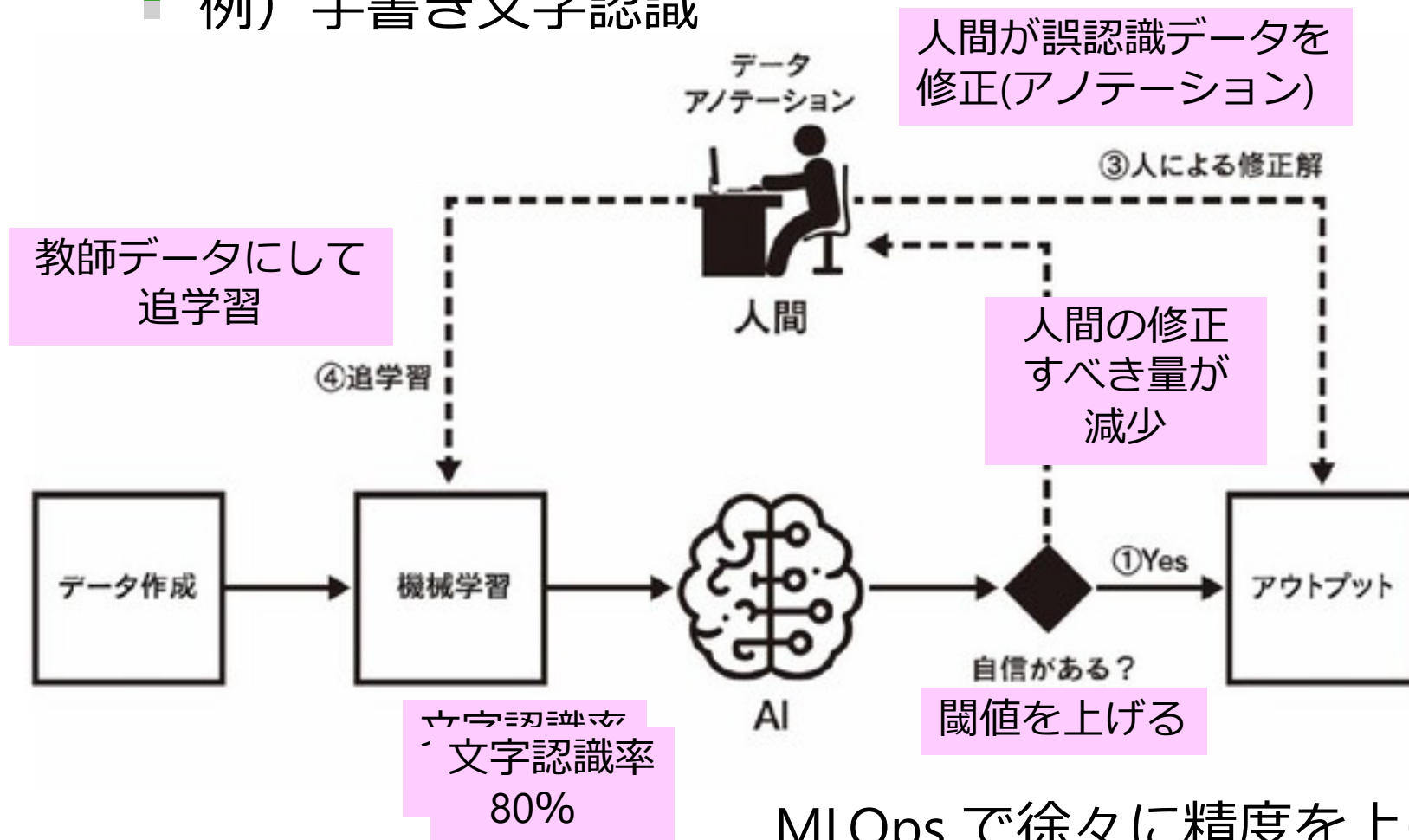
❖ ヒューマン・イン・ザ・ループ (Human-in-the-Loop : HITL)

- AI による自動化・自律化が進んだシステム
 - 実サービスでも一部の判断や制御に人間を介在させる
- AIと協調する新たなシステムへの進化
 - 単調な作業は機械に行わせて人間は高次元の仕事に集中
 - AI では処理できない判断は人間に委ねる



❖ ヒューマン・イン・ザ・ループの構築

■ 例) 手書き文字認識



MLOps で徐々に精度を上げる

❖ 人力検査型

- AI が間違っていないかどうかを人間がチェック
- 人間がすべて検査するよりも手間がかからない
 - 例) 顧客の入力情報の読取

❖ 人間バックアップ型

- AIが上手く対応できない例外処理を人間が対応
- 定型的な対応は AI が実施
 - 例) チャットボットの顧客対応

❖ 監査型

- AI が判断できない微妙な判断を人間が代行
 - 例) クレジットカードの不正利用判定