



そのAI、どこまで"台湾"か？iKala 程世嘉：TMMLU+でソブリンAIの繁体字中国語評価基準を確立

繁体字中国語を「話せる」から、台湾を本当に「理解する」へ。TMMLU+ v1.1が台湾の実情に即した66分野の問題集を更新。次の段階はAIエージェントとマルチモーダルの評価へ。生成AIをめぐる競争は後半戦に入りました。モデル間の性能差が急速に縮まるなか、競争の焦点は「誰が最大のモデルを持つか」から「誰がAIを実際の現場で活用できるか」へと移っています。iKalaの共同創業者兼会長である程世嘉(Sege Cheng)は、台湾がソブリンAI(主権型AI)を議論する際、独自の基盤モデルを構築することだけにとどまるべきではないと考えています。計算資源、モデル、データ、アプリケーションに至るまで、私たちはどれほど自律的に開発・運用する力を持っているのか。この問いを改めて考える必要があるとしています。



iKala共同創業者兼会長の程世嘉(Sege)は、AIをめぐる競争はまだ3回に入っただけだと指摘。台湾がソブリンAIを論じるには、まず自らがどれほどの自律性を確保できているかを考えるべきだとしています。

各国がソブリンAIを競うなか、台湾にとっての鍵とは

程世嘉は、ソブリンAIを「計算資源」「モデル」「データ」「アプリケーション」の4つの要素に分けて捉えています。計算資源はAIを動かせるかどうかを、モデルはその能力の範囲を、データはAIが私たちをどう理解するかを左右します。そしてアプリケーションは、技術が最終的に価値を生み出せるかどうかを決めます。

2023年、台湾の産官学が相次いで基盤モデルの開発に乗り出すなか、程世嘉率いるiKalaは別の問いを出発点にしました。誰もがモデルをつくっているなら、そのモデルの良し悪しは誰が定義するのか、という問いです。

AIの専門分野への特化が進むにつれ、この問いはさらに重要になっています。Gartnerは、2030年には生成AIソリューションの90%が、ドメイン特化型モデル(Domain-Specific Models、DxM)を採用すると予測しています。今後、企業が向き合うのは少数の大規模モデルではなく、業界、タスク、利用場面ごとに構築された数多くのAIです。

程世嘉は、ここにこそソブリンAIの鍵があると考えています。台湾に必要なのは、世界最先端のモデルを利用することに加え、自らのデータと基準を用いて判断する力です。そのモデルは私たちの言語を理解しているか。私たちの産業を理解しているか。求める仕事を最後まで遂行できるか。それを自ら見極める必要があります。

「モデルをつくる」から「物差しをつくる」へ

AIを定義する力を台湾に

こうした問題意識から生まれたのが、TMMLU+です。

市場が繁体字中国語モデルの開発を競うなか、iKalaは「競争に参加するのではなく、まず評価の物差しをつくる」という道を選びました。TMMLU+ v1は66科目にわたる22,690問を収録し、前身のTMMLUの6倍の規模を備えています。研究成果はCOLM 2024で正式に発表され、その後、GoogleやMetaなどの国際的な研究チームによる繁体字中国語モデルの評価にも採用されました。

「オープンソースとして公開したのは、誰もが自由に使えるよう、広く役立ててもらいたいと考えたからです」



TMMLU+は9月に新バージョンv1.1を公開([Hugging Face](#) / [GitHub](#))

程世嘉は、健全なAIエコシステムには、共通して従うルールが必要だと考えています。そのためTMMLU+は、iKalaだけのためのものではありません。産業界と学术界が共同で利

用・検証・改善し、台湾発の評価基準をエコシステム全体で共有する土台にすることを目指しています。

そこで重要になるのが、台湾の実情に根ざした「オリジナルデータ」です。TMMLU+は、インターネット上のコンテンツを単に転載したり、再構成したりしたものではありません。台湾の状況を理解し、内容を抽出し、枠組みを調整するために時間をかけ、継続的にデータの品質を確認しています。モデルが「中国語を話せる」からといって、台湾の法制度、文化的背景、日々の暮らしを本当に理解しているとは限らないためです。

AIが検索や意思決定、情報取得の新たな窓口になるにつれ、世界のAIエコシステムにおける繁体字中国語の相対的な弱さは、単なる言語の問題ではなく、台湾のデータ主権にも関わる問題になっています。信頼できる台湾独自のデータを継続的に蓄積してこそ、台湾は自らの評価基準を築き、台湾に適したAIとは何かを定義できるのです。

「答えられる」から「実行できる」へ

企業が求めるのは現場で使えるAI

AIがチャットからエージェントへと進化するにつれ、評価の問いも「正しく答えられるか」から「仕事を完遂できるか」へと変わっています。

企業が本当に頭を悩ませているのは、使えるモデルがないことではありません。時間と数百万規模の予算を投じてPoC(概念実証)を終えた後で、モデルが繁体字中国語の文脈を理解していない、台湾の規範に適合していない、あるいは公開ベンチマークの高得点と実際の業務フローでの性能に差があると判明することです。そのためiKalaは、評価の対象を大規模言語モデル(LLM)からAIエージェントへと広げています。

次の段階でiKalaが評価するのは、AIが「回答できるか」だけではなく、「仕事を実行できるか」です。意図を理解し、ツールを正しく使い、テキスト・画像・音声などのマルチモーダル情報を処理して、企業から任されたタスクを完遂できるかを検証します。

これにより、評価は技術的なスコアを測るものから、企業の意思決定を支えるものへと変わります。市場が「1つの汎用モデル」から「複数の専門分野に特化したモデル」へ移るなか、企業が本当に必要とするのは、必ずしも世界最高のスコアを持つAIではありません。自社のデータ、利用場面、コスト条件のもとで、最も適したAIです。



[写真キャプション: 原文が「程世嘉將」で途切れているため、iKala側に要確認]

程世嘉は次のように述べています。「私たちは、TMMLU+が台湾においてAIの能力を定義し、産業界の信頼を築くための共通基盤になることを目指しています」。TMMLU+はすでにv1.1 ([Hugging Face](#) / [GitHub](#)) を公開しており、既存の66分野を再点検し、時の経過に伴って見直しが必要となる法律・社会・政策や行政命令に関する設問を更新することで、評価の有効性を維持しています。

世界のモデルが変わり続けるなか、台湾が自らの手に確保しておくべきものは、独自のデータと基準、そしてAIを選び、測り、定義する力です。

【会社概要】

iKala (アイカラ) について

iKalaは、AIとデータをビジネスの中核に組み込み、企業のより迅速でスマートな意思決定を可能にします。データから意思決定へのスムーズな移行を導くことで、企業のAI変革を強力にサポー

ト。独自のファーストパーティデータと、何十億ものグローバルなソーシャルシグナルから構築されたiKalaの高度なインテリジェンスを掛け合わせ、最適なフルネイティブAIソリューションをお届けします。

台湾に本社を構え、世界規模でビジネスを展開するiKalaは、フォーチュン500に名を連ねる大企業をはじめ、190カ国以上の1,000社を超える企業や50,000以上のブランドから信頼を獲得しています。

- **iKala** 公式サイト: <https://ikala.ai>
- **Kolr** 公式サイト: <https://kolr.ai>
- **Kuroma** 公式サイト: <https://kuroma.ai>