

AIの出力ではなく、プロンプトやコンテキストを修正する

ユーザープロンプトとシステムプロンプトがあります。出力がうまくいかない場合、毎回修正するのではなく、ユーザープロンプトにルールを追加するべきです。

以前のエッセイ「AI翻訳出力における名前の問題」で述べたように、中国語の名前には重複が多く存在します。人間でさえ英語の名前を中国語の名前に正確に翻訳できないことがあるため、機械にとってはさらに困難です。

名前の翻訳ルールを定義するのは難しくありません。実際の生活では、文脈が必要になることがあります。例えば、約30人のクラスで、2人が同じ英語名（ピンイン翻訳）を持っている場合、それらの間で翻訳するのは難しくなります。外見などの追加情報が必要になります。

しかし、特定の文脈—例えば誰かのブログ、講義プラットフォーム、ユーザーの連絡先リストなど—では、対象となる名前のリストが大幅に小さくなります。このような場合、AIは完璧に翻訳することができます。

したがって、API呼び出しにおいてこれらのルールを定義すべきです。出力を変更するのではなく、一時的に生成されたコンテンツを調整するのではなく、根本原因を修正することに焦点を当ててください。結果ではなく、その理由を修正することに集中しましょう。

AIコードエディタは、AIチャットボットよりも賢いです。なぜなら、より大きなコンテキストを持っているからです。これにより、出力についてより良い推論を行うことができます。

同じことが画像、音声、ビデオにも当てはまります。これらのツールは、より大きな文脈を提供するべきです。例えば、AI作成ツールに大量のビデオ、音声クリップ、ポッドキャストが与えられれば、それらを基に新しいものをより効果的に生成することができます。

これはRAG（Retrieval-Augmented Generation）と似ています。AI創作ツールにとって、結果が特定されすぎず、また一般的すぎないバランスを取ることが重要です。しかし、結果をより具体的にするためのツールや機能を提供する必要があります。例えば、ChatGPTにはプロジェクト機能があり、ファイルをアップロードしてそれらと対話することができます。他の創作ツールも同様の機能を備える必要があります。