

**LLM などの自然言語処理における課題を解決する独自技術が、
日米に加えて、韓国で特許査定を取得
複数分野のコーパスの特性や構造を維持した解析が可能に**

株式会社 FRONTEO（本社：東京都港区、代表取締役社長：守本 正宏、以下 FRONTEO）は、自社開発 AI エンジン「KIBIT（キビット）」による自然言語処理において、分散表現*¹における特定のコーパス*²の特性や構造を維持したまま複数のコーパスを融合させる技術「Concept Encoder2（コンセプトエンコーダーツー）」について、韓国特許庁より特許査定通知を受けたことのお知らせします。なお、本技術の特許性が認められたのは、日本（2023年11月16日特許登録、特許第7386466号）と米国（2024年7月2日特許登録、特許第12026461号）に続き、第3国目となります。

名称：データ解析装置およびデータ解析プログラム

登録番号：10-2689965

登録日：2024年7月25日

一般的な自然言語処理の過程において、複数の分野のコーパスを統合するとデータが一体化してしまい、それぞれの固有の特性や構造が損なわれることが共通課題とされてきました。本技術により、この共通課題が解決されることで、各分野において、専門性を兼ね備えた自然言語処理 AI 開発および活用を通じた、幅広い事業・ソリューションへの展開が可能になります。

また、同じく当社独自技術である単語および文章の高性能な分散表現を、少ない計算量で算出可能な Concept Encoder（特許番号：日本 6346367、米国 11544309）*³と本技術を組み合わせることで、最適な分散表現を短時間で導き出し、複数のコーパスを横断的かつ効果的に解析することが可能になるため、相乗効果が生まれます。

*1: 単語や文章の意味を次元数の低いベクトルに変換して表す方法

*2: 自然言語研究・解析のために文章を集積した大規模データベース

*3: 自社開発の特化型 AI「KIBIT（キビット）」を構成する自然言語解析技術

■ Concept Encoderについて

「Concept Encoder（商標：conceptencoder）」は、自社開発の特化型AI「KIBIT」のコア技術であり、FRONTEOが独自に開発した自然言語解析技術です。当社がターゲットとする専門性の高いテキスト情報を扱うにあたり、専門用語や低頻度語も含めた膨大な語彙および文章に対して、高性能かつ優れた計算効率でベクトル化することを目的に2018年に開発しました。一般に専門領域のコーパスでは、低頻度語を含むことで語彙数が増加し、計算量が増大しますが、当技術はSub-word化のような分布仮説に従わない手法を避け、分布仮説に忠実に基づき解析的に解くことでこれを実現しています。当社の基幹技術の一つであるこの技術によって新しい関連性（未知情報）のディスカバリを実現します。

また、技術的進歩が極めて早い自然言語処理AI領域においてはBERT(Google、2018年)やLLM等の関連技術が公開されていますが、特にリーガルテック分野やライフサイエンス分野等、専門性の高い分野において、それらの関連技術と比較した際のConcept Encoderの自然言語解析技術をコア技術にもつ特化型AI「KIBIT」の優位性は社内の実証実験によって確認されています。

	KIBIT (Concept Encoder)	大規模言語モデル (LLM)	一般的な embeddingモデル	Transformer ベースモデル
特徴	<ul style="list-style-type: none"> 特許取得：日本、米国(2018) 高性能な分散表現 少ない計算量 文章と単語を同じ空間に埋め込む 解析的に解くため、解は再現性を有する 	<ul style="list-style-type: none"> 例：ChatGPT (Open AI) 10億~数兆のパラメタの最適化計算が必要 最適化によって解くため、解は不安定かつパラメタチューニングのコストが生じる 	<ul style="list-style-type: none"> 例：FastText (META) Word2Vec (Google) 語彙数×ベクトル次元の最適化計算が必要 最適化によって解くため、解は不安定かつパラメタチューニングのコストが生じる 	<ul style="list-style-type: none"> 例：BERT (Google) 1億を超えるパラメタの最適化計算が必要 最適化によって解くため、解は不安定かつパラメタチューニングのコストが生じる
計算	解析的に導出	学習(最適化計算)により導出	学習(最適化計算)により導出	学習(最適化計算)により導出
ベクトル化対象	単語および文章	単語および文章	基本的に単語のみ	単語および文章
分布仮説	○	×	△ Sub-word化等の機能による	×
精度	◎	○ 専門領域に弱い	△	○
計算量	◎	×	○	△

図1：KIBIT（Concept Encoder）のコア技術である分散表現獲得技術と類似技術の比較

※会社名、商品名は、各社の商標または登録商標です。

■ Concept Encoder2について

「Concept Encoder2」は、FRONTEOが独自に開発した自然言語解析技術です。複数のコーパスを横断的に扱う際に単純にコーパスを融合してしまうと、それぞれのコーパスが持つ特性や意味構造が崩れてしまい最適な解析ができないという自然言語処理の共通課題を解決するために2023年に開発しました。当該技術により、例えば複数の学術分野の文献を取り扱う場合や文語体情報と口語体情報を融合して取り扱う場合など、領域横断的に最適な解析を実現します。

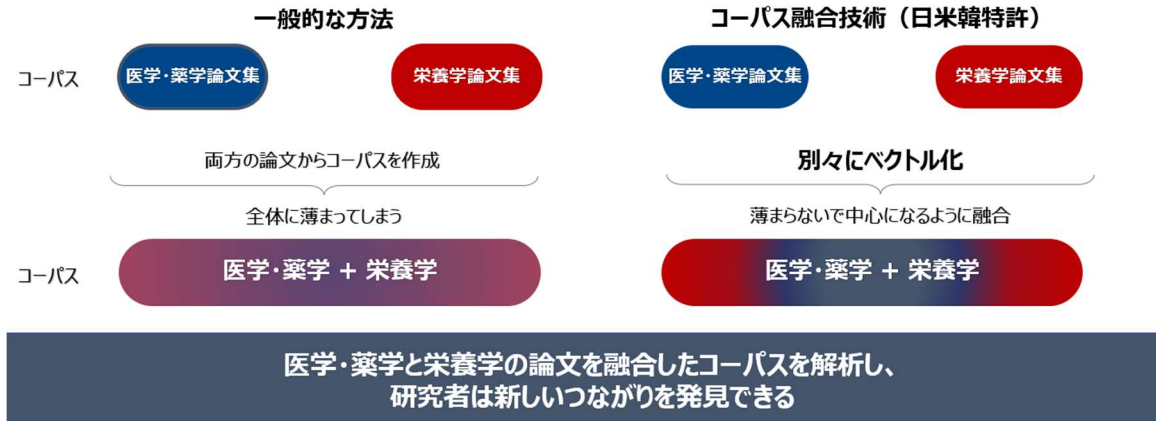


図2：コーパス融合技術のイメージ
(医学・薬学論文集と栄養学論文集を融合する例)

■ FRONTEOについて URL: <https://www.fronteo.com/>

FRONTEOは、自社開発の特化型AI「KIBIT（キビット）」の提供を通じて、日夜、社会課題と向き合う各分野の専門家の判断を支援し、イノベーションの起点を創造しています。当社独自の自然言語処理技術（日米特許取得）は、汎用型AIとは異なり、教師データの量およびコンピューティングパワーに依存することなく、高速かつ高精度での解析を可能にします。加えて、解析した情報をマップ化（構造を可視化）する特許技術を活用することで、「KIBIT」が専門家のインサイトにダイレクトに働きかけることができ、近年、KIBITの技術が創薬の仮説生成や標的探索にも生かされています。

KIBITの独自技術およびアプローチを通じて、「記録に埋もれたリスクとチャンスを見逃さないソリューションを提供し、情報社会のフェアネスを実現する」理念の実現に向けて、ライフサイエンスAI、ビジネスインテリジェンス、経済安全保障、リーガルテックAIの各分野で社会実装を推進しています。

2003年8月創業、2007年6月26日東証マザーズ（現：東証グロース）上場。日本、米国、韓国、台湾で事業を展開。資本金 3,048,772 千円（2024年3月31日時点）。

※FRONTEO、KIBIT、conceptencoderはFRONTEOの日本における登録商標です。

<報道関係者のお問い合わせ先>

株式会社 FRONTEO 広報担当

pr_contact@fronteo.com

<ライフサイエンス AI 事業に関するお問合せ先>

株式会社 FRONTEO ライフサイエンス AI 事業本部

<https://lifescience.fronteo.com/contact>