

FRONTEO ヘルスケア、クラウド対応とデータパッケージを備えた「創薬研究支援 AI システム」を提供開始

1400 万の PubMed 論文、170 万の Open Targets データを備え、最短 1 週間で利用可能。創薬のターゲット候補発見をスピードアップ

株式会社FRONTEO（本社：東京都港区、代表取締役社長：守本正宏）の子会社で、人工知能による医療・介護等の情報解析ソリューションを提供する株式会社FRONTEOヘルスケア（東京都港区、代表取締役社長：西川久仁子）は、ヘルスケア業界向けに独自開発した「客観性」「透明性」「再現性」を特徴として備える人工知能エンジン「Concept Encoder（コンセプトエンコーダー、特許第6346367号）」を用いた「創薬研究支援AIシステム」を本日より提供開始します。

FRONTEOヘルスケアでは、2018年11月にベクトル演算を用いた自然言語処理技術を使った「Concept Encoder」による[新規医薬品探索技術の提供を開始し](#)、現在まで複数の製薬企業に導入されています（※3ページ目の「導入ユーザの反応」を参照）。医薬品開発に従事する研究者が、自らの新しい仮説と公開データベース情報にある文献の記載内容との関連を調べたり、仮説と関連しそうな疾患や同じような遺伝子を変動させる薬との関連を探索し、創薬のターゲット候補の発見に利用されています。

「Concept Encoder」を用いた創薬研究支援AIシステムは、この新規医薬品探索技術をベースに、文献や遺伝子発現などの情報を含むデータベースをあらかじめパッケージ化し、クラウドサービスとして提供することで、より手軽に製薬企業の創薬研究における候補化合物発見のスピードアップを支援するものです。

創薬研究支援AIシステムの主な特徴は、以下のとおりです。

1. 創薬の調査過程で活用が必須となる、MeSH^{*1}タグ付きのPubMed^{*2}論文1,400万と、Open Targets^{*3}のデータ170万をデータベースとして備え、AIエンジン「Concept Encoder」があらかじめデータを学習済みで、即時に探索・解析が可能。
2. 申し込みから最短で約1週間での利用開始が可能（クラウドサービスの場合）。オンプレミスやプライベートクラウドでの導入・利用も可能。
3. ベクトル演算によるAI解析を行うため、辞書登録型のAIと比べ、使用方法が検索ツール並みに容易。
4. 処理が軽いため、オンプレミスの場合、数百万円レベルのサーバー^{*4}で稼働が可能で、スーパーコンピュータや大規模サーバー群などの大型設備が不要。
5. 従来の創薬研究では、仮説から開発承認まで約3-4週間かかる時間を大幅に短縮できる効果が見込まれる。（※次ページ図参照）

製薬企業の医薬品研究開発部門では、創薬研究において、論文情報と公開データベースのチェックを日常的に行い、開発に関連する情報を更新することが不可欠となっています。一方、国内外の医薬研究の動向をタイムリーに追い続けるために、研究者の時間と労力が多大に費やされる点が大きな課題ともなっています。

創薬研究支援AIシステムは、PubMedとOpen Targetsの定期的に更新された最新のデータベース情報と論文のテキスト情報を学習させてあるため、研究者が自らの仮説を自然文で作成し、入力するだけで、仮説に関連する最新の論文やターゲット遺伝子ネットワーク等との関連性の強さがスコア（数値）で可視化されます。またカスタマイズにより、製薬企業が過去保管してきた文書データを読み込ませて、活用することも可能です。

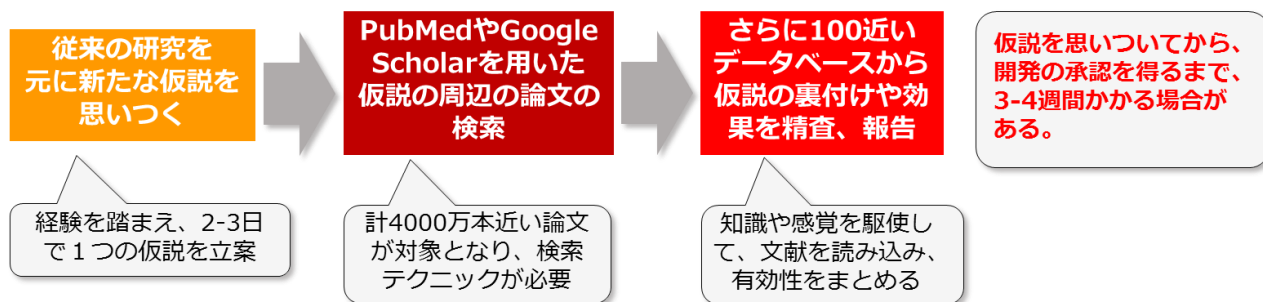
従来の論文解析で用いられてきた辞書型や文法学習型のAIは、使用するために研究者がコーディングの知識を学ぶ必要があったり、常に単語と意味のデータベースのアップデートが必要でした。さらに初期のトライアルやPoC（概念実証）には、何ヶ月もの期間と大規模な予算が必要と言われていました。

創薬研究支援AIシステムは、人工知能エンジンConcept Encoderの特徴であるベクトル演算を用いることで、単語や文書同士の関連や類似度を数値化できるため、辞書では定義しにくい曖昧な概念もベクトル化で関係を表すことができ、データベースの頻繁なアップデートが不要です。また、解析には、近似式を用いるため、解析やデータのアップデートにはマシンパワーが軽く済むことも使いやすい点です。まるで検索ツールのように、自然文を入力し、例えば、「ガンは除く」といった不要な要素を自在に引き去り、差分を抽出の対象とした解析が可能な、取扱いやすさも特徴の1つです。

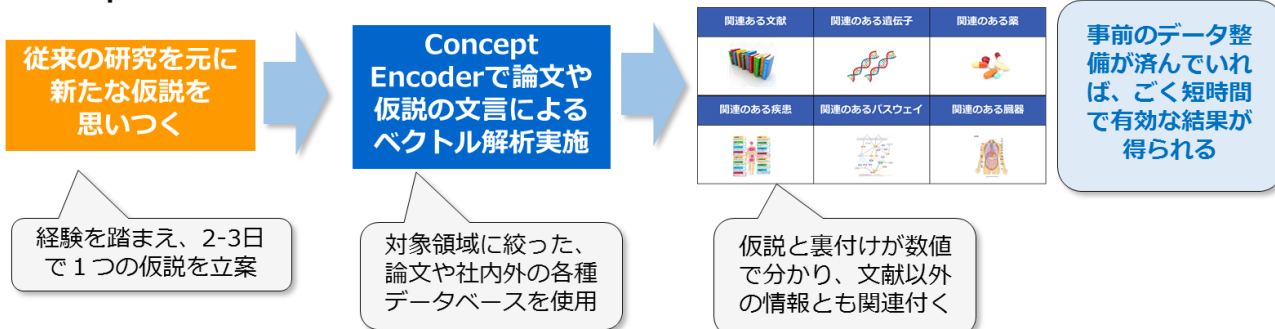
FRONTEO ヘルスケアでは、今後さらに Concept Encoder を軸とした情報解析ソリューションの提供を広げ、ヘルスケア業界全体の発展に貢献してまいります。

【従来の論文情報検索と、Concept Encoder を用いた創薬研究支援 AI システムの違い】


【従来】



【Concept Encoder を用いた創薬研究支援AIシステム】



【創薬研究支援 AI システム 検索画面】



concept Encoder
VC Report

Overview | clinical | animal model | molecular mechanism | Full View

Report

- レポートに関連する遺伝子を抽出
- 仮説などを入れる
- 関連の深いトップ100論文概要
- Animalモデル, clinical, molecular mechanismに自動分類
- 関連度スコアでソート
- グループ毎の論文数の推移
- 動物モデルと臨床とのギャップ

Related Genes

query	entrez_id	symbol	name	class
ADP3	360	ADP3	equiplorin 3 IgM blood group	molecular_mechar
CCL2	6347	CCL2	C-C motif chemokine ligand 2	molecular_mechar
IL1B	3553	IL1B	interleukin 1 beta	molecular_mechar
IL6	3569	IL6	interleukin 6	molecular_mechar
INS	3630	INS	insulin	molecular_mechar

Number of publications

Legend: animal_model (blue), clinical (orange), molecular_mechanism (green)

top 100 pubmed articles

id	score	class	abst
28472477	0.0117	clinical	Context: Contracting data have be... Fibronectin Type III Domai
18513333	0.0101	animal_model	AIM: The aim of this study was to... Bezafibrate prevents hepat
27321932	0.0535	animal_model	Cyclooxygenase-2 (COX-2) is invo... Cyclooxygenase-2 expres

Animal and Clinical pathway similarity score

38.5%

【論文の探索研究に Concept Encoder を導入した製薬企業の反応】

～Concept Encoderが優れている点～

① 解析結果のオリジナリティ

“様々なテキストマイニングツールを調べたが、アウトプットが単語同士を繋ぐネットワーク図のようなものが多かった。概念図を出されても先に進めないが、Concept Encoderはスコアが一覧で出てくるため、使い勝手が良い”

② 遺伝子情報等とリンク

“複数の情報源をリンクさせて解析できるのが有益。ただの論文検索とは異なり、解析結果に広がりが生まれる”

③ 対象データの拡張性

“探索の対象をフレキシブルに変えられるため、既存のデータベースを活かしたり、将来的に他のシステムとの連携が必要になっても容易に拡張できる”



- ※1 MeSH : Medical Subject Headings の略。PubMed の論文で医学を主題とした内容であることを表わす。
- ※2 PubMed : 生物や医学の論文を検索できるデータベース。NLM (米国国立医学図書館) 内にあるNCBI (国立生物科学情報センター) が作成。
- ※3 Open Targets : 官民のパートナーシップによる創薬ターゲット発見のためデータベース
- ※4 創薬研究支援 AI システムが稼動可能なサーバーのスペック例 ⇒ CPU 数 : 8 コア、メモリ : 256GB、ストレージ容量 : 6TB

【Concept Encoder について】 URL: <https://www.fronteo-healthcare.com/conceptencoder>

Concept Encoder (コンセプトエンコーダー) は、FRONTEO ヘルスケアがヘルスケア・インダストリーに特化して開発した人工知能 (AI) です。自由記述のテキストデータを大量に含むヘルスケア関連のビッグデータを、エビデンス (根拠) に基づいて有効に解析・活用することを目標に開発しました。ヘルスケア従事者の共通認識である「エビデンスに基づいた医療 (EBM)」に欠かせない有意差検定などの統計学的手法を自然言語解析に導入、実現しています。Concept Encoder はテキスト以外のデータとの共解析も可能であり、ヘルスケア領域に蓄積されてきた遺伝子発現情報・バイタルや各種検査値などの「数値データ」との共解析の研究を進めています。特許登録番号 : 特許第 6 3 4 6 3 6 7 号

【FRONTEO ヘルスケア 会社概要】 URL: <http://www.fronteo-healthcare.com/>

名称 : 株式会社 FRONTEO ヘルスケア
設立 : 2015 年 4 月 16 日
資本金 : 327,000 千円 (資本準備金 210,000 千円含まず)
代表者 : 代表取締役社長 西川 久仁子
事業内容 : 診断支援、ヘルスケア業務支援、製薬業界支援、その他医療・介護等の情報解析に関する事業

【FRONTEO について】 URL: <http://www.fronteo.com/>

株式会社FRONTEOは、独自開発の人工知能エンジン「KIBIT」により、ビッグデータなどの情報解析を支援するデータ解析企業です。国際訴訟などに必要な電子データの証拠保全と調査・分析を行うeディスカバリ (電子証拠開示) や、デジタルフォレンジック調査を支援する企業として2003年8月に設立。自社開発のデータ解析プラットフォーム「Lit i View (リット・アイ・ビュー)」、日・中・韓・英の複数言語に対応した「Predictive Coding (プレディクティブ・コーディング)」技術などを駆使し、企業に訴訟対策支援を提供しています。このリーガルテックAI事業で培われ、発展した「KIBIT」を始めとする独自の人工知能関連技術は、専門家の経験や勘などの「暗黙知」を学び、人の思考の解析から、未来の行動の予測を実現します。ビジネスインテリジェンスやヘルスケアなどの領域でAIソリューション事業を展開し、FinTechやRegTechに加え、「働き方改革」でも実績をあげています。2007年6月26日東証マザーズ、2013年5月16日NASDAQ上場。資本金2,559,206千円 (2019年3月31日現在)。2016年7月1日付けで株式会社UBICより現在の社名に変更しております。

〈ヘルスケア案件に関するお問合せ先〉

株式会社 FRONTEO ヘルスケア

TEL 03-5463-6330 FAX 03-5463-7578 Email: inquiry@fronteo-healthcare.com

〈報道関係者のお問合せ先〉

株式会社 FRONTEO 広報担当 池内

TEL: 03-5463-6380 FAX: 03-5463-6345 Email: pr_contact@fronteo.com