# KNOWLEDGEABLE AI: 知識を効率的に追加・編集・蓄積できるAIを作る

山田 育矢（Studio Ousia, 東北大学, 名古屋大学, 理化学研究所） ikuya@ikuya.net ikuyamada

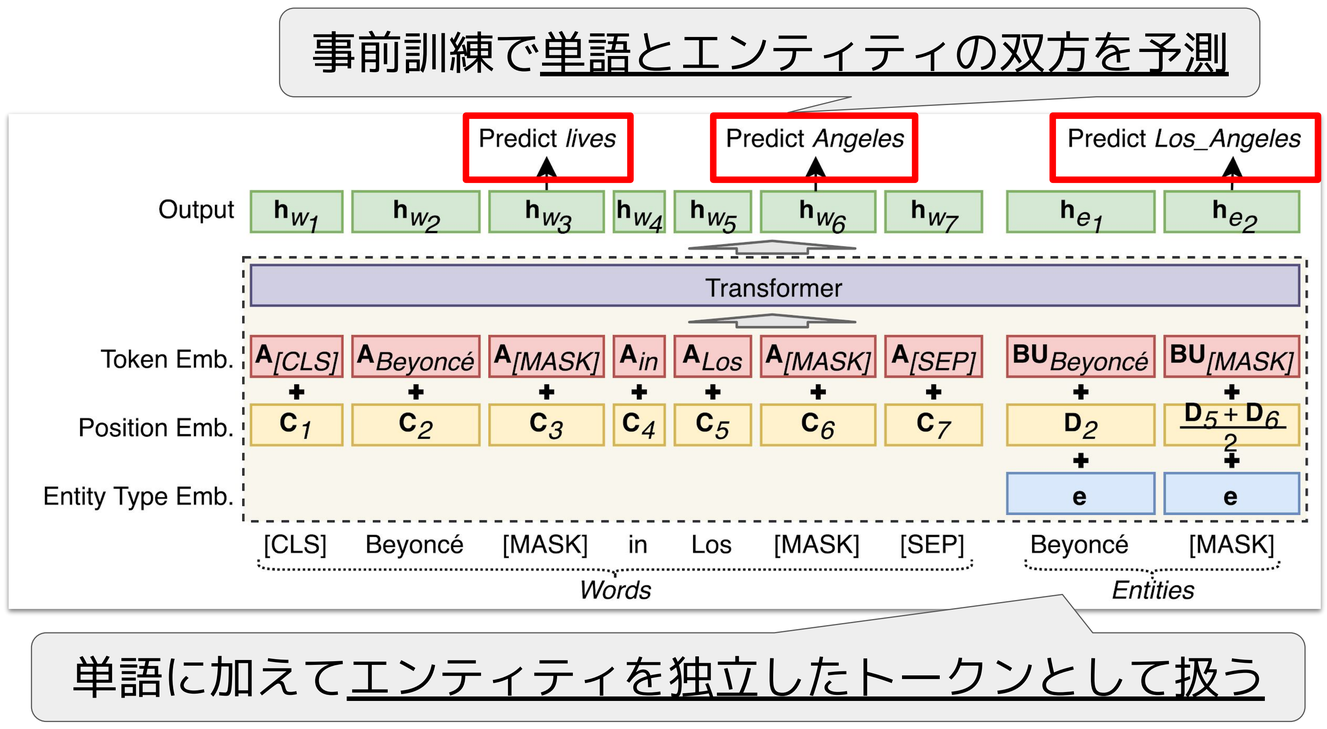
## 言語モデルの知識の扱いの改善に興味を持っています

- **組織・集団でのみ通用するローカルな知識**の扱い
- **コーパスに頻出しない専門的な知識**の扱い
- 世界の変化に伴う**知識の動的な更新**
- **低資源言語での知識の活用**の促進

## LUKE: 言語モデルでのエンティティの扱いの改善 (2020-)

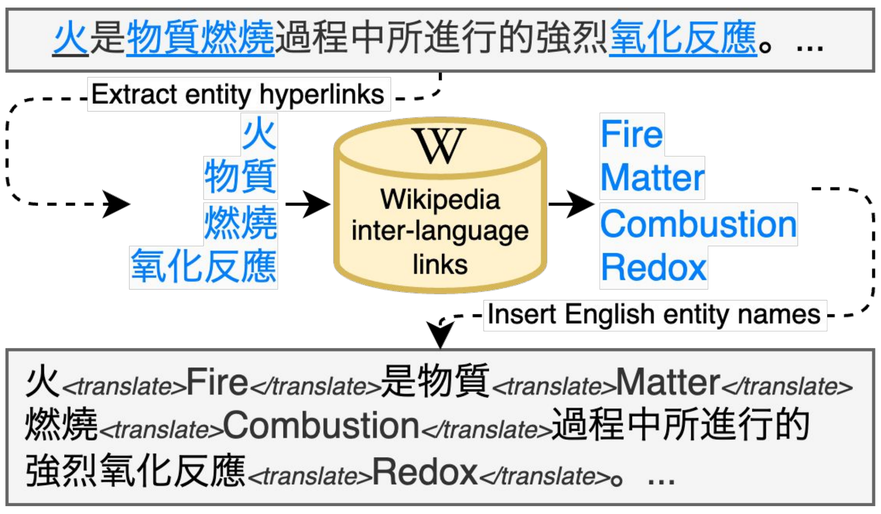
**エンティティを独立したトークンとして明示的に扱うことで様々な応用が可能に**

- 単語（サブワード）に加え、エンティティを独立した入力／出力トークンとして扱うエンコーダモデル
- モデルはエンティティの語彙と埋め込みを持つ
- **SOTA性能:** エンティティを扱う複数の情報抽出・質問応答のタスクでSOTAを獲得 (EMNLP 2020)
- **多言語化:** 言語を横断したエンティティ埋込みを通じた言語間転移の促進 (ACL 2022)
- **知識の追加・更新:** エンティティの説明文からの埋込みの動的な補完 (EMNLP FIndings2022)
- **ドメイン特化:** 企業の知識を豊富に保持した経済情報特化モデル (ユーザベース社との共同研究; 2024)




**LUKEのア ーキテクチャとWikipediaでの事前訓練**

## LEIA: 英語から他の低資源な言語への知識転移の促進 (ACL Findings 2024)




**Wikipediaを使ったLEIAのデータ拡張**

**言語モデルは主に英語コーパスで訓練されているため**
**英語で学習済みの知識が、他の言語では有効に使えないことがある**

**Wikipediaを使った単純なデータ拡張による言語間知識転移の促進**

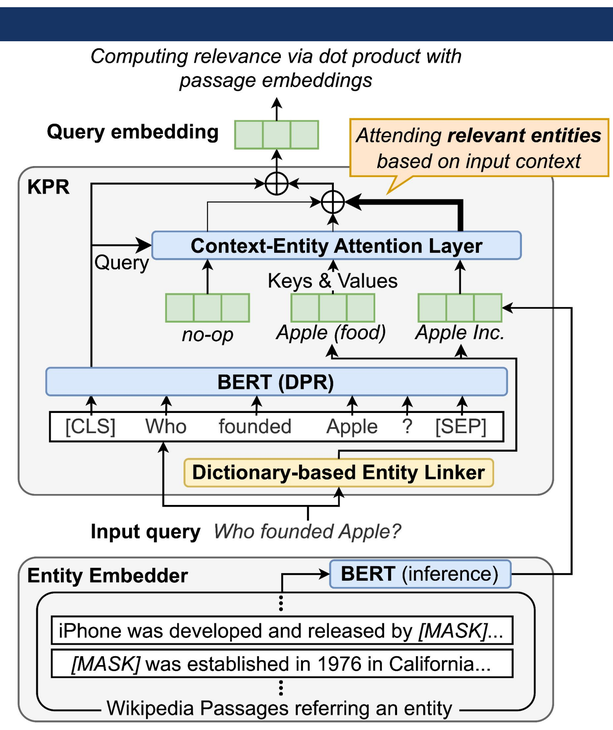
- **手法:** Wikipediaテキストのリンクの隣に該当する英語名を挿入して訓練
- **効果:** 訓練時に英語名に紐づいた知識が想起されて他の言語に転移
- **結果:**日本語、中国語、アラビア語、ヒンディー語、スワヒリ語などの7言語の多言語質問応答（X-CSQA、X-CODAH）で顕著な性能改善

## KPR: 稀な知識が必要な検索の改善 (EMNLP Findings 2025)

**稀なエンティティ（人名・会社名等）を含むクエリは、埋め込み検索の性能が顕著に低いため、RAGで有効な知識をLLMに入力できない**

**エンティティの知識を動的に注入可能にすることによる検索の大幅な性能改善**

- **手法:** エンティティの知識を取りいれる単純なアテンション機構の導入と言及テキストからのエンティティ埋め込みの言語モデル推論による動的生成
- **結果:** 稀なエンティティのクエリを多く含むデータセット（EntityQuestions）で検索の性能が劇的に改善（recall@20 56.8% -> 69.4%）




**KPRのアーキテクチャ**

## 現在興味があること

- エンティティ埋込みを活用した多言語埋め込みモデル
- 画像からのエンティティ知識のVision LLMへの導入
- 解釈できるエンティティ埋め込みを通じた知識の注入
- より良く知識を取り込めるDeep Researchモデル