

# 留学報告書

安永迪弘  
2022年12月

2019年9月より Stanford大学にてコンピュータサイエンス(CS)の博士学生をしております安永迪弘と申します。今回は4年目夏～秋学期の生活を振り返りたいと思います。

## 1. 研究

早いものでPhD 4年目になり、job talkやdissertation (博士論文) を見据えて研究を積み上げていくよう、指導教授からアドバイスをいただく時期になりました。

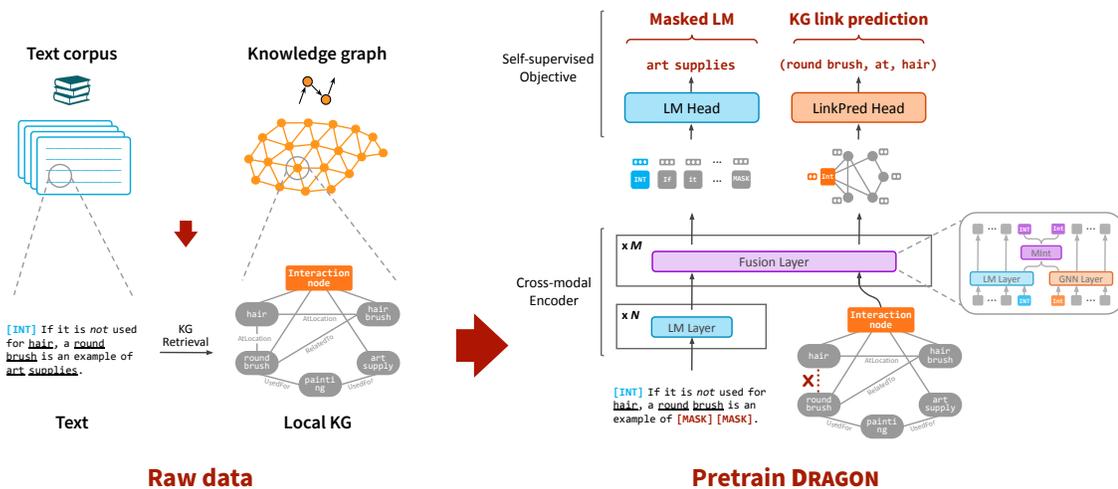
PhD課程の研究テーマとして、知識と推論を駆使する自然言語処理の技術を開発することを目指しています。Googleなどの検索エンジン、Siriなどの質問応答システムを始め、多くの自然言語処理タスクは解答を出すために関連する背景知識を集め、それに基づいて推論をする必要があります。このようなシステムを開発する上での大きな課題は、背景知識となる情報源が様々な形式 (テキストに加え、知識グラフ、画像、データベースなど) に散らばっており、それらを統合する方法が必要であることです。方法の開発・改善に向け、この半年間、主に次の2つの研究課題に取り組みました。

### (1) テキストと知識グラフの大規模な統合を行うには？

これまで自身の研究 ([QAGNN](#), [GreaseLM](#)) にてテキストと知識グラフを統合する方法を一步步開発してきましたが、モデルの学習は、人の手によってラベル付された小規模な質問応答データを用いて行ってきました。今回はそれを一般化し、**ラベル付されていないテキストデータ**と知識グラフを用いて学習できる方法「[DRAGON](#)」の開発に取り組みました。具体的には、テキストにおける「言語モデリング」と知識グラフにおける「リンク予測」という二つのラベルなし学習方法 (self-supervised learning) を組み合わせています。最近の機械学習全般として、Web上に大量にあるラベル付されていない (= 安価な) データを使って上手くモデルを学習できないか、という研究が注目されています。ラベルなし学習の強みとして、多量・多様なデータからモデルを学習でき、その結果、幅広いタスク<sup>1</sup> にモデルを適応できるようになります。

---

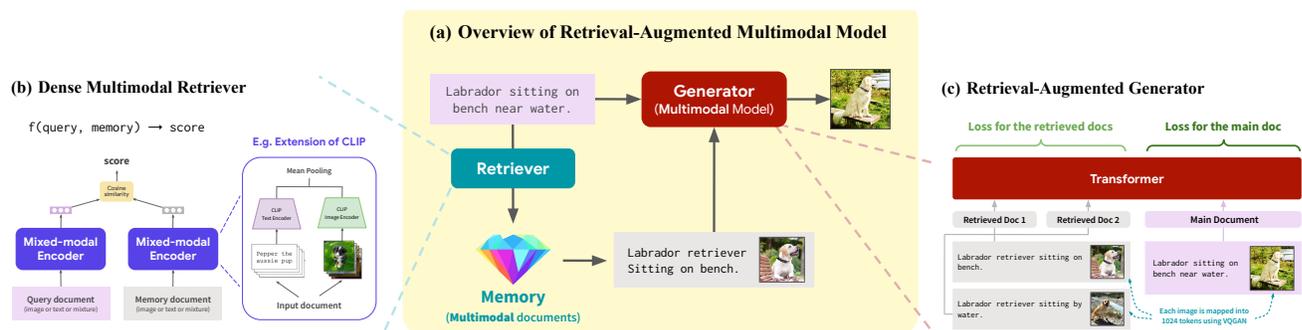
<sup>1</sup> 例えば、様々な分野の質問応答タスク



論文： M. Yasunaga, A. Bosselut, H. Ren, X. Zhang, C.D. Manning, P. Liang\*, J. Leskovec\*. "DRAGON: Deep Bidirectional Language-Knowledge Graph Pretraining". NeurIPS 2022. (<https://arxiv.org/abs/2210.09338>)

## (2) テキストと画像を統合するには？

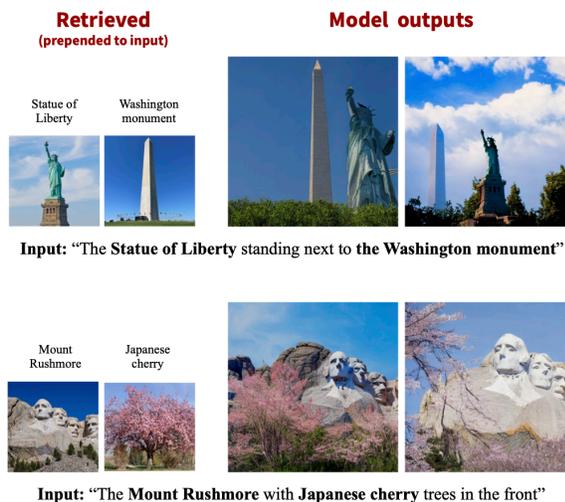
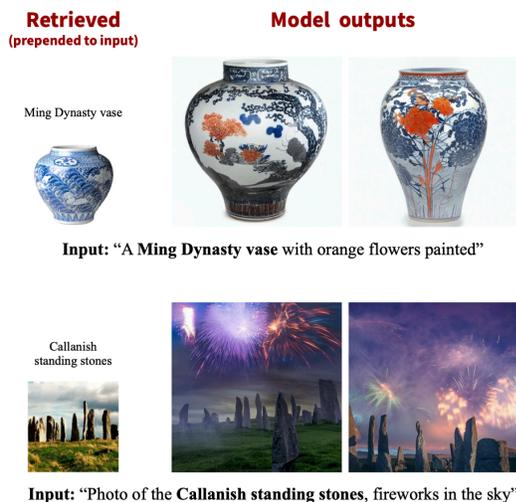
最近、[BERT](#) や [GPT-3](#) などの言語モデルが持つ知識<sup>2</sup> を拡張するため、Wikipediaなど外部のテキストを随時検索・参照しながら言語生成できるようにする Retrieval-Augmented Language Models (検索拡張言語モデル) の研究<sup>3</sup> が盛んになっています。ところが、Web全体を参照可能な情報源として考えてみれば、テキストだけでなく画像も有用かつ大部分を占める情報源です。そのため、テキストと画像の両方を検索・参照・そして生成できる、[マルチモーダルな検索拡張言語モデル](#)の開発に取り組みました。



<sup>2</sup> 例えば、「アメリカの首都はワシントンDCである」など

<sup>3</sup> 例えば、Metaの[DPR/RAG](#), Googleの[REALM](#), DeepMindの[RETRO](#)など

このモデルの面白い結果として、細かい背景知識を必要とするテキスト及び画像の生成を高精度で行うことができるようになります (例えば、下図の「明朝時代の陶磁器」)。また、[GPT-3](#) や [DALL-E](#) など近年のテキスト・画像生成モデルは数十億以上のパラメーターで大量の学習をすることが必要でしたが、検索拡張機能を与えて知識を外部参照化することで、少量の計算資源・小さくて効率良いモデルでも同等以上の生成精度を達成できることが分かり、とてもエキサイティングな研究になりました。



この研究は主にMeta AIでのインターンプロジェクトとして取り組みました。Meta AI は検索拡張言語モデルの研究を牽引している研究所の一つで、検索拡張機能の実装<sup>4</sup> について非常に勉強になりました。また、ありがたいことに、大学にはなかった規模の計算資源を使った実験ができ、貴重な経験になりました。

論文： M. Yasunaga, A. Aghajanyan, W. Shi, R. James, J. Leskovec, P. Liang, M. Lewis, L. Zettlemoyer, W. Yih. "Retrieval-Augmented Multimodal Language Modeling". (<https://arxiv.org/abs/2211.12561>)

<sup>4</sup> MetaAIから検索ライブラリがオープンソースされています: <https://github.com/facebookresearch/faiss>

## 2. 最後に

初めて企業研究所でインターンをしたということもあり、インターンが非常に印象に残る経験になりました。とりわけ、メンターとチームに恵まれたことに感謝しています。特に、メンターの Scott (Wen-tau) 先生と Luke 先生は、インターンプロジェクトが企業のみならず自分の博士論文内容にも寄与できるようプロジェクト方針を一緒に考えてくれ、また検索拡張機能やマルチモーダルモデルに詳しい周りの研究者も熱心に協力してくれ、とても充実した研究を行うことができました。そして、チームの研究者層が厚く、知見が広いのも企業研究所の一つの魅力だと感じました。例えば、自分のようなPhDインターンから、日々研究・実装しているフルタイム研究者とエンジニア、そしてシニアな研究マネージャー・大学と企業両方の所属を持つ教授まで、幅広い研究者が一つのチームになってプロジェクトに取り組むスタイルは大学にはあまりなかった面白い経験です。

得られた経験を活かして、引き続き良い研究をしていこうと思っています。船井財団にはいつもサポートしていただき、本当に感謝しております。



Poo Poo Point Trail



Mount Rainier

インターンの3ヶ月間、シアトルで夏を過ごしました。シアトルは自然が豊かでとても好きになりました。様々なハイキングコースがあっておすすめです (写真)。また交通の便が良い・アジア料理へのアクセスが良いところも好きです (なんと鼎泰豊が複数軒あります)。