

教員名	吉永 直樹	研究場所	生産技術研究所 (生研)	研究分野	自然言語処理・計算言語学
-----	-------	------	--------------	------	--------------

# ことばを正しく、高速に「計算」する

人間は言語を基に思考し、社会での経験を記録して他者と共有することができます。本研究室では、人の様々な言語活動を代替・支援するとともに、人が紡ぐ無数の言葉からその心と社会の動きを読み解くことを目指し、言葉を正しく、高速に「計算」する工学的研究（**自然言語処理**）に取り組んでいます。最速・最小・最高精度のモデルの追求が、ときに確率的に、またときに規則正しく振る舞う自然現象としての言語の「かたち」に迫る理学的研究（**計算言語学**）につながり、最終的には人間の知能の解明やその改善にも寄与すると考えています。

<https://www.tkl.iis.u-tokyo.ac.jp/~ynaga/>

本研究室では、言語・文化・学問の背景の異なる学生が独自の視点で行う研究こそが、多様な言語の本質を理解する上で重要と考えており、学生は、主体的かつ自由に研究テーマを設定することが期待されています。定型的で退屈な方法論にとらわれず、新たに研究課題を生み出す意欲のある方の参加を歓迎します。

## 自然言語処理: 言語の運用能力を実現する

ウェブなど大量のテキストデータをもとに膨大な計算資源を用いて学習された大規模言語モデルは、言語の指示に応じて多様なタスクを実行できることから、人間の幅広い知的活動を代替しつつあります。しかしながら、現状の言語モデルには、実世界知識の記憶・想起や算術・論理演算といった離散的な現象を連続的に扱うことに起因する幻覚や、学習データのバイアスの増幅といった課題が残されています。

本研究室では、工学的視点から、言語モデルの動作原理を解明し、モデルの本質的な改善と応用の開拓を進めています。具体的には、実世界知識に基づくモデルの分析・圧縮・分解・合成、推論時計算や検索拡張生成、画像・音声を含む、マルチモーダル多言語処理、社会情報学への応用に取り組んでいます。

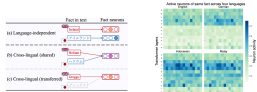
## 計算言語学: 言語の数理的構造を探究する

大規模言語モデルは膨大な連続パラメータを最適化することで言語の振る舞いを高精度に近似する一方で解釈性に課題があり、言語という自然現象の数理的理解は、依然として発展途上にあります。また、日々変化する実世界知識や、誤りが許されない算術演算など規則的汎化を要する記号処理においては、学習データの枯渇やコストの観点から、単純な規模拡大による精度向上にも限界があります。

本研究室では、理学的視点から、言語モデルが近似する記号処理の構造的改善や利活用を進めています。具体的には、アルゴリズムやデータ構造など計算機科学的手法を駆使して記号処理の性能限界の突破を図り、記号処理機構を創発する学習戦略を設計しています。現在は、言語モデルが捉える未知の構造的知識を見出すことで、人知を拡張することを目指しています。

### 大規模言語モデルの内部機序の分析

- 機械論的解釈可能性 (知識想起・算術演算)
- 幻覚の自己認知



大規模言語モデルの知性の解明

### 大規模言語モデルの高度化

- 語彙の最適化に基づく最小 LLM の探索
- 検索を用いた生成の効率化
- 推論時計算・検索拡張生成



言語モデルの性能限界の拡張

### マルチモーダル・多言語処理

- 画像・音声付き多言語資源の構築
- 視覚・音声言語モデルに基づく生成・評価



言語による人の認知の理解

### 大規模言語モデルの応用の探求

- ユーザの模倣と個人適応
- SNSにおける情報変容の解明



Explore diverse LLM applications

### 経験主義的記号処理

- 機械学習モデルの記号処理への再解釈

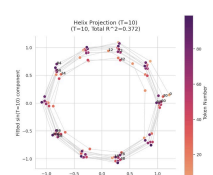
意味のない人がいる。  
shumi no nai hito ga iru .

Pattern	Word	POS (level 1)
意味	意味	NOUN
の	の	ADP
な	な	ADJ
人	人	NOUN
が	が	ADP
いる	いる	VERB
。	。	PUNCT

基礎解析の効率の限界を20年ぶりに更新 (M2 MacBook Air で 100万文/秒)

### 言語モデルの記号処理機構の解明

- 算術処理を可能とする幾何的構造の解明と誘導



合理主義的連続処理の実現

本研究室は豊田・合田研究室と密に連携しており、研究環境（データ、計算機、学生部屋）を共用しています。合同ミーティングを通じ大局的視点からフィードバックを受けられるほか、全量マイクロブログ投稿と実世界情報を組み合わせた社会分析など**自然言語処理の枠に留まらないスケール感のある学際的研究**も可能です。

質問や見学の希望などありましたら、気軽に吉永 (ynaga@iis.u-tokyo.ac.jp) までご連絡ください。