

# 自然言語処理：言葉がわかる人工知能

## ～体験しよう！人工知能の最前線～



電気情報物理工学科 情報工学コース

### 乾研究室

情報伝達の最も重要なメディアは、日本語や英語といった、人間が日常で使う言語＝**自然言語**です。

本研究室では、自然言語で表現・伝達・蓄積される情報をコンピュータで処理するための基礎理論・基盤技術・応用技術の研究を行っています。

例えば、楽しい会話ができる対話ロボット、英語論文の執筆を支援するシステム、記述式試験の採点をサポートするシステムなどを開発しています。



### 知識を使った多様な応答が可能な雑談対話システム

深層学習による発話生成とルールによる応答を組み合わせ、知識に基づく対話を行います。

仙台のデザートにハマってます

ずんだ餅、美味しいですね！

デザートといえばアフォガートって知ってますか？

仙台といえば…ここでクイズです！

### 自然言語生成技術を用いたライティング支援システム

作りたい文章の内容に適した語彙や言い回しを自動で提案し、文章作成・推敲を支援します。

As mentioned previously, we focus the task of correcting grammar errors automatically

こんな感じの文をもっと自然に書きたいなあ…

earlier や before はどうでしょう？

As mentioned previously, we focus the task of correcting grammar errors automatically | on を追加しましょう

automatically correcting grammatical errors という語順のほうが自然です

### 国語記述式問題の自動採点システム

国語の記述式問題の答案を、複数の採点基準と照らし合わせて自動で採点します。

答案

日本人のように言外に含みを持たせるのではなく、**他者に理解してもらうために言葉を尽くす「対決」のスタンスが西洋特有の文化を導いてきた。**

**9点**

- 採点基準1: 西洋、西洋人…2点
- 採点基準2: 他人は自分と違う…3点、対決のスタンス…2点
- 採点基準3: 同意を得るために…3点
- 採点基準4: 言葉を尽くして…3点、他人を説得する…3点
- 採点基準5: 文末が「こと」でない…-1点

### 47都道府県の方言の自動翻訳

最先端の自動翻訳の技術を応用して、共通語と全国の方言の間の翻訳を行います。

雨が冷たい！

雨がしゃっこい！

雨がひゃっこい！

雨がひやい！

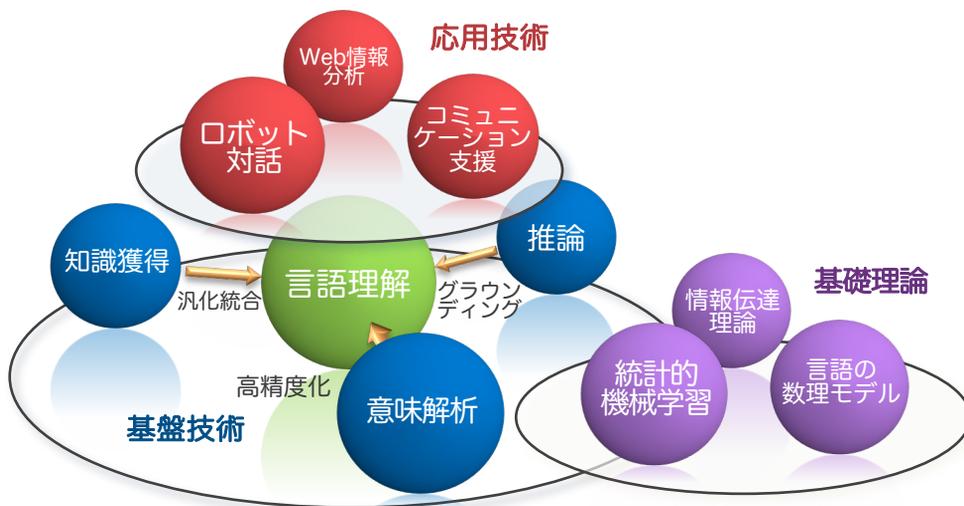
雨がひじゅるー！

# 乾研究室

<https://www.nlp.ecei.tohoku.ac.jp>

情報伝達のための最も重要なメディアは、日本語や英語などの、人間が日常で使っている**ことば言語**です。このような言語を、情報科学ではプログラミング言語などの形式言語と区別して「**自然言語**」と呼びます。

私たちは、自然言語で表現・伝達・蓄積される情報をコンピュータで処理するための基礎理論・基盤技術・応用技術を研究しています。**自然言語処理** (natural language processing) ・ **人工知能** (artificial intelligence) ・ **機械学習** (machine learning) ・ **知識情報処理** (knowledge processing) ・ **計算言語学** (computational linguistics) ・ **コミュニケーション科学** (communication science) などと呼ばれる領域が私たちのフィールドです。



インターネットやウェブの爆発的な普及によって、誰でも**大量の情報**を入手・蓄積・発信できる時代になりました。しかしその一方で、あまりにも多くの情報がネット上に無秩序に分散しているために、欲しい情報をうまく探せない、重要な情報の存在に気づかない、情報が信用できるかどうか分からない、といった問題が日常的に起こっています。

さて、ここでいう情報はその多くが自然言語ですから、求められるのは**自然言語処理**です。膨大な言語情報をコンピュータで自動的に収集・選別・分析できれば、自動的に他言語へ翻訳することや、対話的に人に伝えることも可能となり、我々を取り巻く情報環境は大きく変わるでしょう。こうした自然言語処理による**Web情報分析**や**コミュニケーション支援**、**ロボット対話**といった応用技術の重要性が急速に高まっています。

これらの技術を実現するためには、究極的には人の**言葉を理解**するコンピュータを開発する必要があります。もちろん、これは簡単なことではありません。しかし、自然言語処理の技術は着実に進歩しています。例えば、これまでコンピュータに決定的に欠けていた**常識的知識**をコンピュータ自身が大量の言語データから自動的に獲得して、より高度な**意味解析**や**推論**に使う、ということが少しずつ可能になり、大きなブレークスルーの兆しが見え始めています。

本研究室では、言葉が分かるとはどういうことか？ コミュニケーションの成立条件は何か？といった情報伝達の仕組みを解明しモデル化する理論的研究、および人間の知的な情報伝達・情報分析を支援するソフトウェアを構築する工学的研究を展開します。**言葉から人の知に迫る**。心躍る研究領域がここにあります。

言語コミュニケーションの仕組みや不思議さに惹かれる人、次世代のWeb情報サービスを企業と組んで作りたい人、数理統計的なモデル化や人工知能の研究をやってみたい人、広く歓迎します。

事前の専門知識は不要です。研究室の中で基礎から勉強します。大切なのは新しいことに挑戦していける好奇心とそれを持続できる根気です。

元気な研究室を一緒に作っていきませんか？