

ことばを速く、正しく「計算」する技術で、人のこころと社会の動きを読み解く

人は言語を用いて考え、社会での体験を記録し、他人に伝えることができます。本研究室では、膨大な言葉から人の心と社会の動きを読み解き、人の情報伝達・理解を支援することを旨として、計算機を用いて言葉を速く、正しく「計算」する工学的研究（自然言語処理）に幅広く取り組んでいます。これらの技術の追求が、ときに確率的に、ときに規則正しく振る舞う自然現象である言語の「かたち」に迫る理学的研究（計算言語学）や人間の知能の働きの解明に繋がると考えています。

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

言葉から人のこころを理解する

直接観測できない人の心を言葉から読み解く研究を行っています。具体的に、テキストから書き手である大衆の価値観を捉える手法を開発しました。また、深層学習を用いて個人が用いる単語の意味の揺らぎの数理的なモデル化を行いました（図1）。現在は、対話システムとの会話を通して話者の心身の状態を把握・改善するプロジェクトに取り組んでいます。

言葉から社会の動きを読み解く

実世界の情報を言語で発信する人々をソーシャルセンサとみなし、膨大な投稿から社会の動きを把握する研究を進めています。具体的に、言語情報の冗長性に着目しテキストの量の変化に応じて言語解析（構文・意味解析）を加速する機械学習手法を考案しました。現在、テキストから新事物などの世界知識を即時検出する知識獲得手法やその集積技術を研究しています（図2）。

人の情報伝達・理解を支援する

人の情報伝達を支援する機械翻訳や対話システムの研究を行っています。これまでに、聞き手の感情を揺さぶる人間的な対話システムや、時間情報や画像など非言語情報を活用したマルチモーダル翻訳・対話システムの研究を行っています（図1）。また言語の壁を越えた汎言語的な多言語処理モデルの開発や文法誤り訂正、テキスト平易化など周辺技術の研究も進めています。

本研究室では、研究を社会実装につなげることを重視しています。これまで、日々蓄積されるソーシャルビッグデータ（図3）を実際に解析して社会の動きを捉えるシステム（図4）を構築してきました。現在は、スマートスピーカに実装した対話システムを用いて、高齢者のヘルスケア支援を進めています。我々は、既存の研究の後追いではなく、新たな研究課題を生み出す意欲のある方の参加を歓迎します。また、漠然とことばを科学したい方の参加も歓迎します。

言葉を通した人のこころの理解と情報伝達の促進

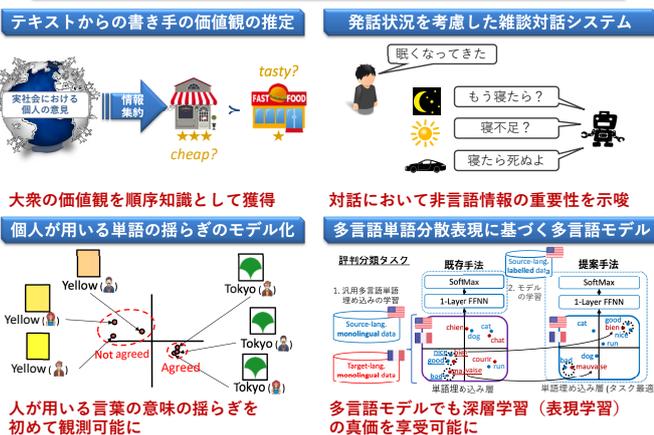


図1: 言葉を通した人のこころの理解と情報伝達の促進

超高速かつ適応的にテキストを読み解く基礎言語解析

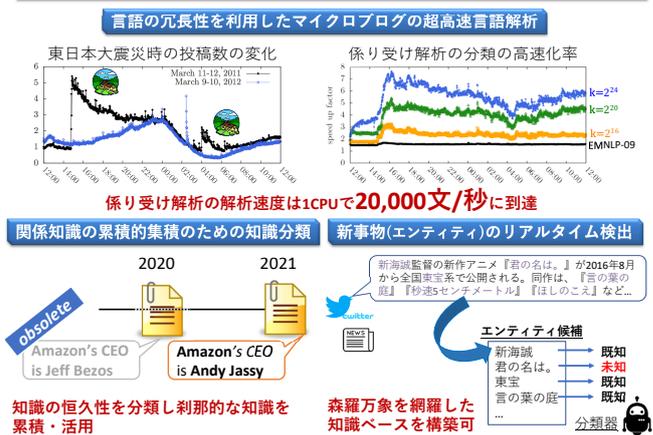


図2: 超高速かつ適応的にテキストを読み解く基礎言語解析

ソーシャル・ビッグデータ



図3: ソーシャルビッグデータ

Figure 4 details: The table shows 'Real-Time Parsing' results for the Great East Japan Earthquake. It lists dates from 2011/03/11 14:30 to 2011/03/11 21:30. The columns include 'ツイート数' (Tweet count), '地震' (Earthquake), '津波' (Tsunami), '原発' (Nuclear power), '避難' (Evacuation), '救助' (Rescue), '被害' (Casualties), '死者' (Deaths), '行方不明' (Missing), '救助隊' (Rescue teams), 'ボランティア' (Volunteers), '寄付' (Donations), '支援' (Support), '情報' (Information), '伝言' (Messages), '連絡' (Contacts), '捜索' (Search), '発見' (Discovery), '救助' (Rescue), '治療' (Treatment), '回復' (Recovery), '帰国' (Return to home), '帰郷' (Return to hometown), '帰省' (Return to hometown), '帰国' (Return to home), '帰郷' (Return to hometown), '帰省' (Return to hometown).

図4: 東日本大震災発生日の話題追跡