

テキストメディア分野 [知-12]

(学術情報メディアセンター 大規模テキストアーカイブ研究分野)

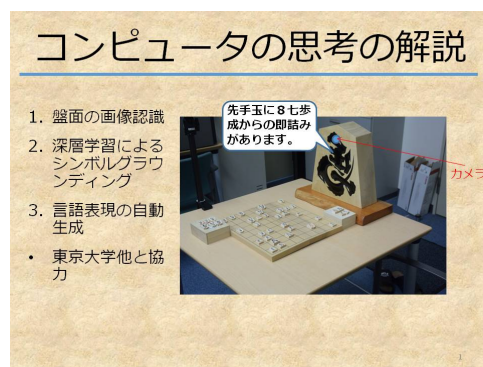
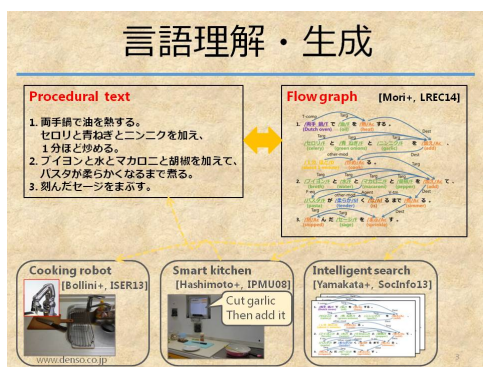
教授 森 信介 forest@i.kyoto-u.ac.jp

助教 亀甲 博貴 kameko@i.kyoto-u.ac.jp

<http://www.lsta.media.kyoto-u.ac.jp/>

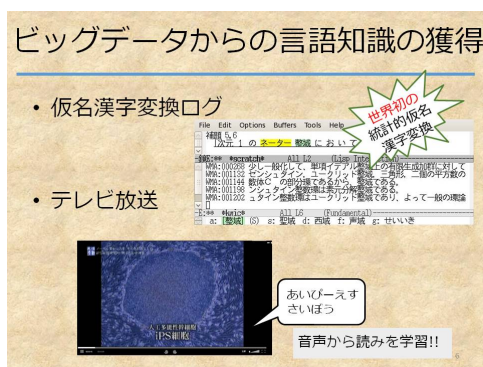


古来、人類の知は文書に記録されてきました。本分野では、これを理解し新たな知を記述できるコンピューターの創造を目指しています。そのための基盤として、自然言語の理解とそのために必要となる基礎的な自然言語処理の研究を行っています。加えて、データ分析や未来予測などのコンピューターの思考、あるいは映像や音声などの他のメディアを言葉で説明する自然言語生成について研究を行っています。



自然言語理解は人工知能の究極の目標の1つです。レシピを代表とする手順書をターゲットとして自然言語理解に取り組んでいます。また、(調理) 映像などからの手順書(レシピ)の自動生成にも取り組んでいます。

ゲームを題材にコンピューターに自身の思考を言葉で説明させる方法の研究をしています。機械学習がよりブラックボックス化する中で、ホットな研究テーマです。証券市場の予測とその解説にも対象を広げています。



ウェブテキスト、テレビ放送、言語処理ツールの利用ログなどから言語知識を獲得し、言語処理に活用することに取り組んでいます。ツイッターなどの言語解析では最高水準のツールを持っています。

形態素解析(単語分割・品詞推定)や係り受け解析などの基礎的な言語処理について研究し、論文発表に留まらずツールの公開も行っています。他の研究テーマで共通に用いられる重要な研究課題です。