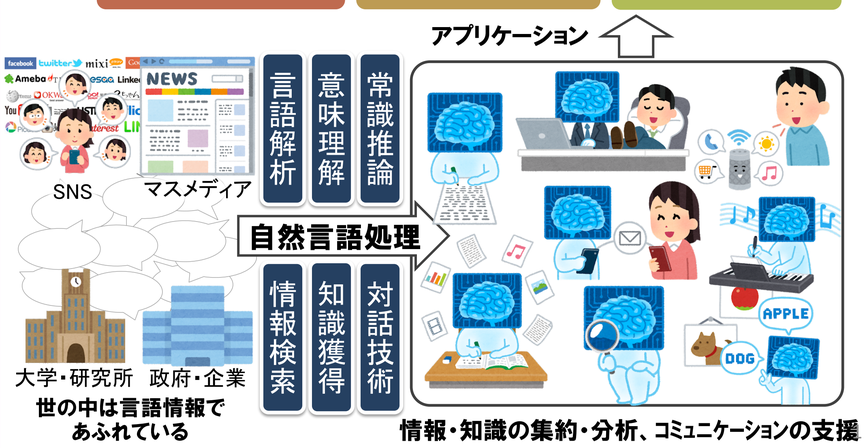
自然言語処理：言葉がわかるAIをつくる

自然言語処理学研究室　　**教授**：乾健太郎、**准教授**：鈴木 潤、**助教**：井之上直也
研究員：Paul Reisert、大内啓樹、Benjamin Heinzerling、三田雅人、Ana Brassard、松田耕史、清野舜、佐々木翔大、塙一晃、高橋容市、山口健史、磯部順子、菅原真由美

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、誰もが日常で使っている言語（自然言語）です。本研究室では、自然言語で表現され、伝達され、蓄積される情報や人の知識をコンピュータで処理するための基礎理論、基盤技術、応用技術に関する研究を行っています。

自然言語処理：応用の場はあらゆる産業へ

科学技術 / 金融・経営 / 自動運転 / 教 育 / 医療・健康 / 製造プロセス / 公共サービス



東北大NLP×産学・学際連携

東北大学 自然言語処理学研究室
メンバーの受賞歴

2019年	Google AI Focused Research Award
2019年	ICLR2019 Outstanding Reviewer Award
2019年	言語処理学会 第25回年次大会 優秀賞
2019年	言語処理学会 第25回年次大会 言語資源賞
2019年	言語処理学会 年次大会 若手奨励賞×2
2019年	言語処理学会 年次大会 最優秀ポスター賞
2019年	東北大学 総長賞・工学部長賞・他
2018年	対話システムライブコンペティション 優秀賞
2018年	人工知能学会 全国大会 優秀賞×2
2018年	人工知能学会 2017年度論文賞
2018年	5th CTJW Best Poster Presentation Award
2018年	情報処理学会 第236回SIG-NL 優秀研究賞
2018年	船井情報科学振興財団 研究奨励賞
2018年	東北大学 総長賞・工学部長賞・他
2018年	言語処理学会 第24回年次大会 優秀賞
2018年	言語処理学会 年次大会 若手奨励賞×3
2018年	トーキン科学技術賞
2017年	情報処理学会 DBS研究会 学生奨励賞
2017年	人工知能学会 SLUD研究会 若手奨励賞
2017年	NLP若手の会 第12回シンポジウム 奨励賞
2017年	人工知能学会 全国大会 優秀賞
2017年	情報処理学会 第231回SIG-NL 学生奨励賞
2017年	EACL2017 Outstanding Paper Award
2017年	2016年度マイクロソフト情報学研究賞
2017年	言語処理学会 第23回年次大会 優秀賞
2017年	東北大学 総長賞・研究科長賞・工学部長賞
2016年	PACLIC-30 Best Paper Honorable Mention
2016年	MIRU2016 優秀賞
2016年	文部科学大臣表彰 若手科学者賞
2016年	IBM faculty Award
2016年	言語処理学会 第22回年次大会 最優秀賞
2016年	言語処理学会 第22回年次大会 優秀賞
2016年	情報処理学会 第225回SIG-NL 優秀研究賞
2016年	東北大学 工学部長賞
2015年	PACLIC-2015 Best Paper Award
2015年	ドコモ・モバイル・サイエンス賞
2015年	NLP若手の会 第10回シンポジウム デモ賞
2015年	東北大学 総長賞
2015年	言語処理学会 2014年度最優秀論文賞
2015年	言語処理学会 第21回年次大会 最優秀賞
2015年	人工知能学会 第29回全国大会 学生奨励賞
2015年	情報処理学会 第221回SIG-NL 優秀研究賞
2014年	ICOAI-2014 Excellent Oral Presentation
2014年	日本学術振興会 科学研究費審査委員表彰
2014年	石田貴記念財団 研究奨励賞
2014年	情報処理学会 2014年度山下記念研究賞
2014年	AMT-2014 Best Paper Award
2014年	言語処理学会 20周年記念論文賞
2014年	東北大学 総長賞
2014年	言語処理学会 2013年度 優秀論文賞
2014年	言語処理学会 第20回年次大会 最優秀賞
2014年	言語処理学会 第20回年次大会 優秀賞
2014年	言語処理学会 第20回年次大会 若手奨励賞
2014年	第3回サイエンス・インカレ JST理事長賞
2014年	第4回WI2研究会 優秀研究賞
2013年	日本リスク研究学会 優秀発表論文賞
2013年	言語処理学会 第19回年次大会 優秀賞
2013年	東北大学 情報科学研究科長賞
2012年	情報処理学会 2012年度 山下記念研究賞
2012年	言語処理学会 第18回年次大会 優秀賞
2012年	情報処理学会 NL研究会 奨励賞
2011年	NLP若手の会 第6回シンポジウム 優秀発表賞
2011年	日本経済新聞社「技術トレンド調査」総合3位
2011年	CICLing-2011 Best Paper Award (First Place)
2010年	言語処理学会 第16回年次大会 最優秀発表賞
2010年	言語処理学会 第16回年次大会 若手奨励賞

「行間を読む」言語理解へ

「行間を読む」

週末に母とやなぎに行ってみました。
うどん屋の名前
うどんを食べるために
私は評判のきつねうどんを頼んだの
私はうどんを食べた
やなぎで
ですが、期待通りの味に大満足。
うどんが美味しいと
きつねうどんの

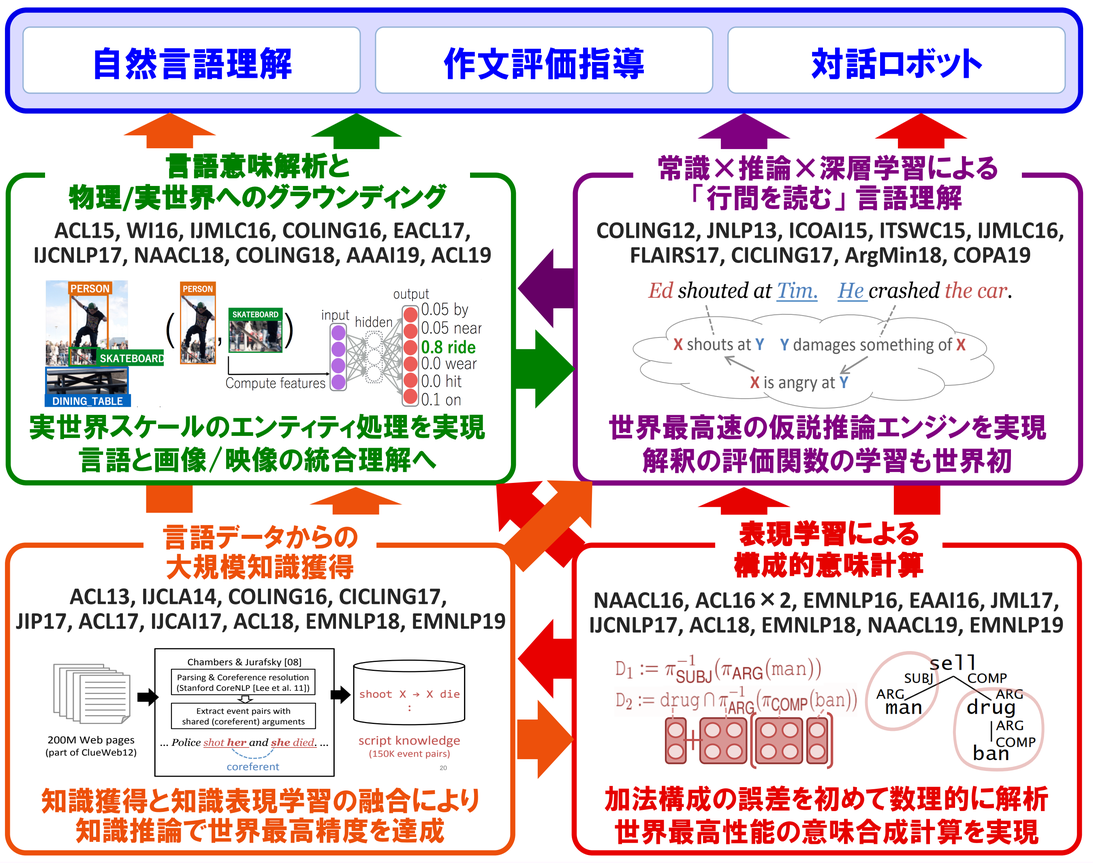
常識もビッグデータから集める

天気が良いので、洗濯物がよく乾く
天気のいい日は洗濯物がよく乾く
最近，天気がいい日が多いので、洗濯物がよく乾く
天気が良くないので、洗濯物が乾くのもままならず
天気がいいから洗濯物がすぐに乾いて
天気が良い日でも洗濯物がなかなか乾かなくて

⇩

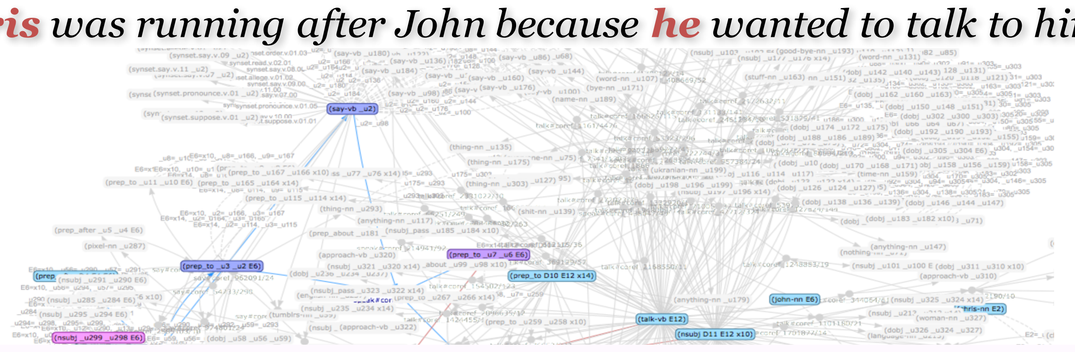
天気が良い（原因）⇒ 洗濯物がよく乾く（結果）

常識推論×深層学習による高度言語理解と対話/教育への応用



東北大で世界最高速の仮説推論システムを開発
（米大システムの1万倍）

Chris was running after John because he wanted to talk to him.



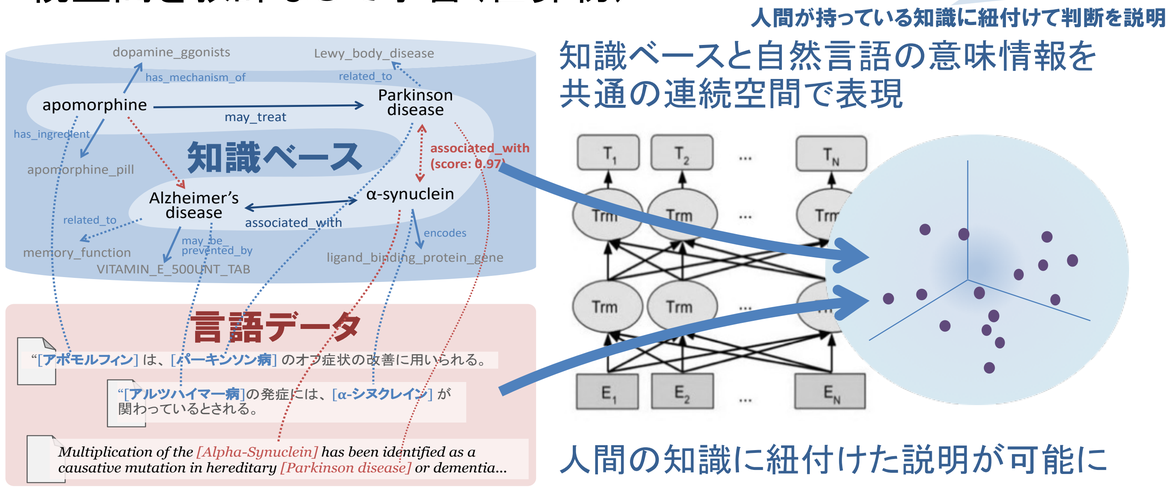
AIシステムの処理結果の解釈性、説明性

［我々のアプローチ］

- 人間が持っている常識や専門知識を深層学習器に教え込み、処理結果を人間の知識に紐付けて説明する仕組みを作る
- 知識上の推論と言語の意味の対応が付くように、両者を共通に表現できる連続空間を教師なしで学習（世界初）

アポモルフィンは**アルツハイマー病**に効果がある可能性がある。**なぜなら、アポモルフィン**は**パーキンソン病**の改善に用いられており（KB）、**パーキンソン病**の発症には**α-シヌクレイン**が関わっている（論文）と…

人間が持っている知識に紐付けて判断を説明



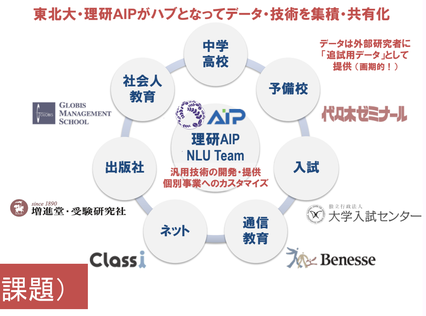
知識ベースと自然言語の意味情報を共通の連続空間で表現

人間の知識に紐付けた説明が可能に

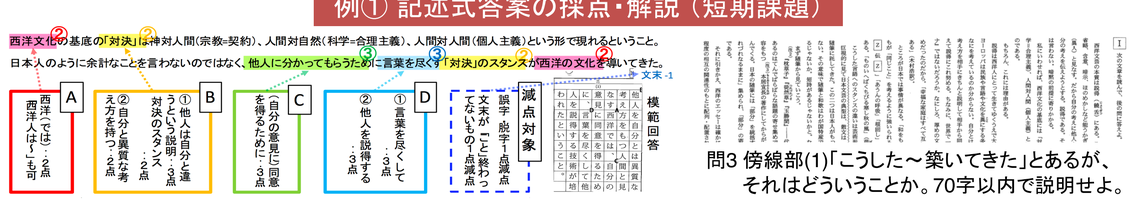
自然言語アセスメント：AI版「赤ペン先生」をつくる

人間の作文・論述（思考）を解析し、「どこがどう良いか/不十分か、どう直すとよいか」を説明できるAI技術

- 社会的ニーズ：思考力・判断力・表現力の育成は国家的課題
  育成には採点・添削・個別指導を要するが、教員リソースは逼迫
- 技術的課題：不完全な文章の解析、文脈からの「意図」の推測等が必要 → 自然言語処理の未解決重要課題（骨のある課題）



例① 記述式答案の採点・解説（短期課題）

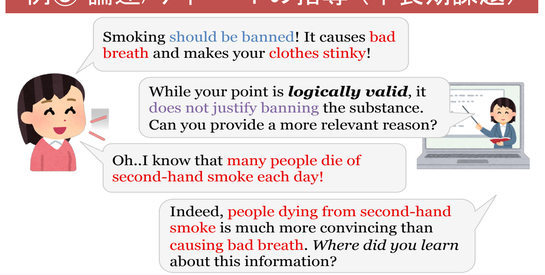


例② 英作文の添削・解説（短中期課題）

トピック：ガーデニング
学習者の解答:
I had ever been to gardening. Because my mother's hobby is gardening. And I beyond gardening club for 11 years ...

＜過去完了＞は、ある過去の時点を明示した上で、それまでの行為・状態を表す場合に用いられます。単に「(これまでに)…したことがある」という経験を表す場合は、＜現在完了＞で表します。
<Past perfect> is used with a certain point of time to talk about something that happened before that point. Use <present perfect> to talk about experiences.

≪have been to≫は「…に行ったことがある」を表します。「ガーデニングをする」という表現で用いられる＜動詞＞を調べてみましょう。
<<have been to>> is used to express ``an experience in visiting'' Use another verb for ``gardening''.

≪beyond≫は＜前置詞＞で「…の向こうで；…を越えて」です。ここは＜動詞＞belongを用いて「…に所属している」と表してみましょう。
<<beyond>> is a < preposition> meaning ``further away'' or ``outside a limit''. Use the <verb> ``belong'' instead.

例③ 論述/ディベートの指導（中長期課題）