

言語で説明できるAIへの道のり

～ 質問応答、自然言語アセスメントを例題として ～

乾 健太郎

東北大学 情報科学研究科
理化学研究所 AIPセンター

乾 健太郎

専門分野：自然言語処理

1995年 東京工業大学 情報理工学研究科 博士修了 博士(工学)

1995年 東京工業大学 情報理工学研究科 助手

1998年 九州工業大学 情報工学部 知能情報工学科 助教授

2002年 奈良先端科学技術大学院大学 情報科学研究科 助教授

2004年 英国Sussex大学 客員研究員(1年間)

2010年 東北大学大学院情報科学研究科 教授

2016年 理化学研究所 革新知能統合研究(AIP)センター
自然言語理解チームリーダー 兼任

言語処理学会 副会長

日本学術会議 連携会員

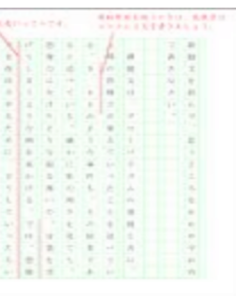
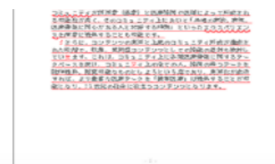
国立情報学研究所 客員教授

NPO法人 ファクトチェック・イニシアティブ 理事

理研AIPセンター 自然言語理解チーム

テーマ：自然言語アセスメント

AI版「赤ペン先生」で教育に革新を！



人間の作文・論述を解析し、「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技術をつくる



理研記述問題採点データセット

2020年7月公開



大学共同利用機関法人 情報・システム研究機構
国立情報学研究所
National Institute of Informatics



情報学研究データリポジトリ

English

HOME データ一覧 研究成果一覧 ユーザフォーラム 組織 関連リンク 問い合わせ

> HOME > データ一覧 > 理研記述問題採点データセット

民間企業提供データ

- Yahoo!データセット
- 楽天データセット
- ニコニコデータセット
- リクルートデータセット
- クックパッドデータセット
- LIFULL HOME'Sデータセット
- 不満調査データセット
- Sansanデータセット
- インテージデータセット
- オリコンデータセット

理研記述問題採点データセット

国立情報学研究所が研究者提供データセットとして理化学研究所から受け入れたデータセットです。

2020/07/14 作成

データ概要

大学受験予備校や教育関連会社にて行われたテストの記述問題の解答に対し、理化学研究所において採点アノテーションを付与したデータセットです。今後も順次同様のデータの追加を予定しています。

1. 代々木ゼミナールデータ（2020年4月時点、9問各約2000件）

2014年～2015年に行われた高校生対象模擬試験の国語記述問題に採点アノテーションを付与したものです。全体点および採点項目ごとの部分点や、その根拠となる単語・フレーズがアノテーションされています。

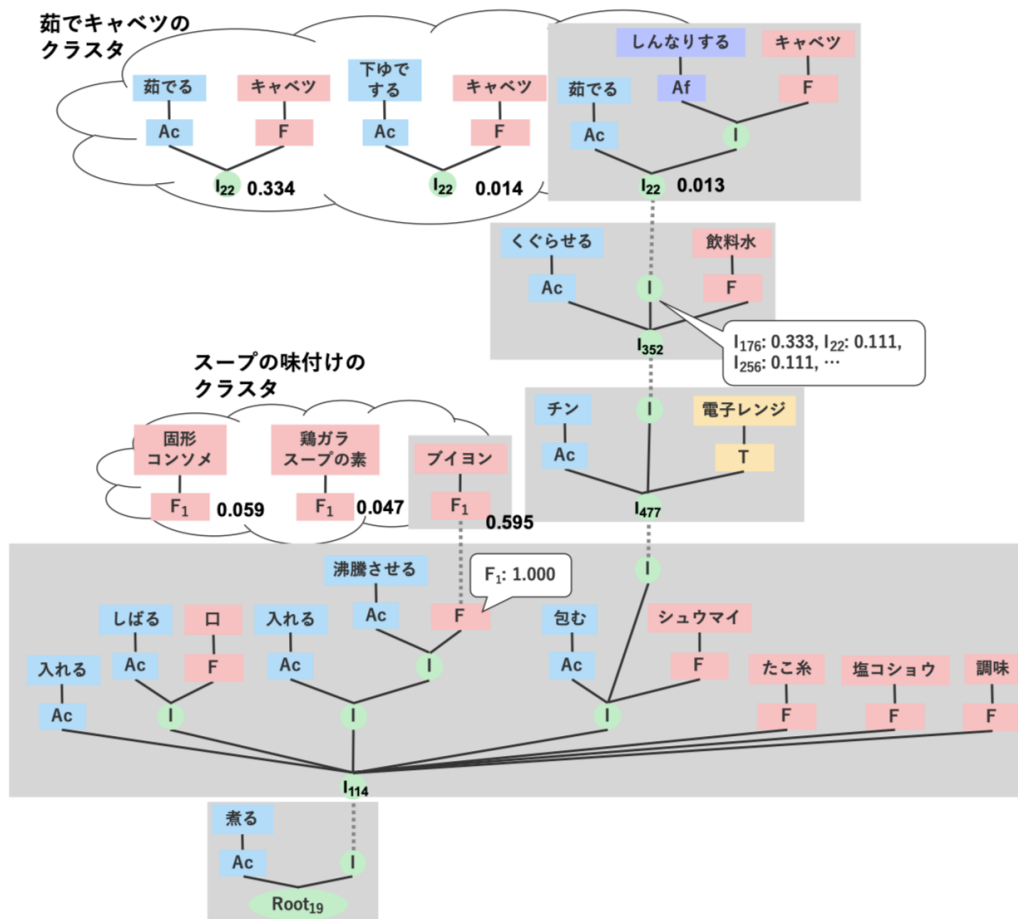
データの詳細は[理化学研究所の「記述答案採点の自動化に関する研究」のページ](#)もご参照ください。

また本データを用いた研究成果として理化学研究所が発表した[国際会議ACL併催ワークショップBEA14の論文](#)もご参照ください。

「クックパッドデータセット」を使って

確率木置換文法と分散表現に基づくレシピ構造木の生成モデルの学習

(吉成+ NLP 2019)



言語で説明できるAIへの道のり

～ 質問応答、自然言語アセスメントを例題として ～

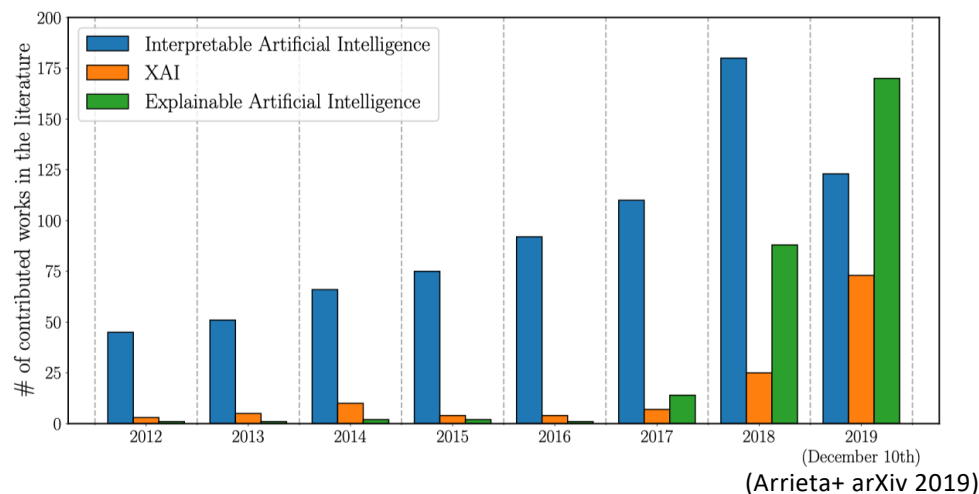
講演内容

1. イントロ: Explainable AI (XAI)
2. DNNと記号推論の融合を通して
3. マイルストーン: タスク設計とデータ構築

Explainable AI (XAI)

「なぜそう判断したか」を説明できる仕組みがほしい

Evolution of the number of publications



一般ユーザ この訳は本当にあって
いるのかな？

正しい訳という
確証が欲しい

翻訳システム
の例

ユーザ : 通路側の席をお願いします
システム : eono klkino, ioa.

開発者

なんかまずい応答して
いるけど大丈夫かな？

デバッグ必要だけど、
どこを修正すれば
いいのだろうか？

対話システム
の例

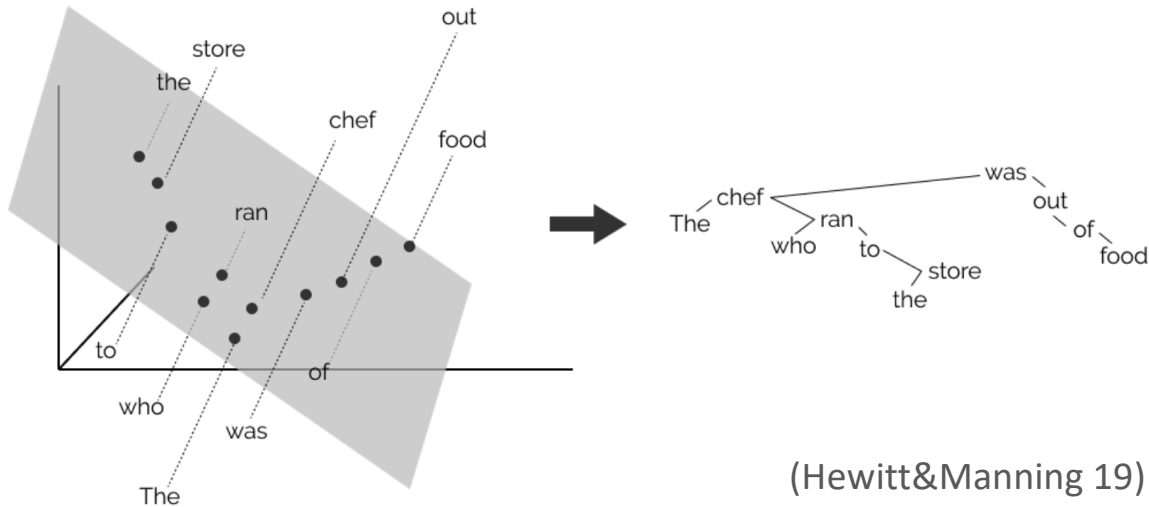
ユーザ : 今日は気が重いから休みたいなあ
システム : ハッピーになれるいいクスリありますよ

Explainable AI (XAI)

- 「何をもって説明とするか」は百家争鳴状態
(原聡: 私のブックマーク「説明可能AI」, 人工知能学会誌, 2019)
- **大域的な説明 (モデル説明):**
 - シンプルなモデルに還元する
 - gradient, attentionなどモデルの中身を見る
- **局所的な説明 (出力説明):**
 - 入力のどこが効いているかを見せる (attention, LIME, etc.)
 - 判別に寄与した訓練事例を見せる
- **説明生成の学習:**
 - 説明生成そのものをend2endに学習
- **説明可能なモデルの設計:**
 - 最初から可読性の高い解釈可能なモデルを作る
(rule lists, rule sets, etc.)

Explainable AI (XAI)

- 大域的な説明（モデル説明）：
 - シンプルなモデルに還元する
 - gradient, attentionなどモデルの中身を見る



Explainable AI (XAI)

- 局所的な説明（出力説明）：
 - 入力のどこが効いているかを見せる（attention, LIME, etc.）
 - 判別に寄与した訓練事例を見せる

入力のどこを見て予測したか

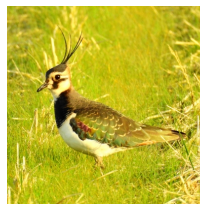


判断：ネコ



猫に注視してネコと予測しているので判断は正しそう
特に耳の部分に特徴があるっぽい

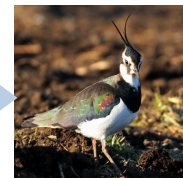
予測の根拠となる訓練事例を提示



タゲリ

予測ラベル

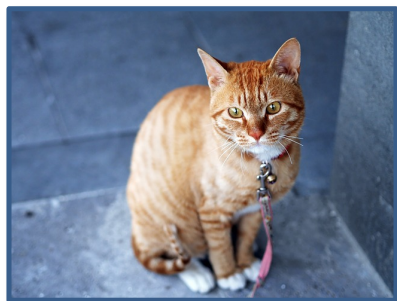
訓練データ



根拠となる
訓練事例

なっとく感？

物体認識



判断：ネコ



猫に注視してネコと予測しているので判断は正しそう
特に耳の部分に特徴があるっぽい

いい感じに新しく知見を得たように
聞こえる。なるほど感あり。

感情分析

判断：ポジティブ

Review 1:

この映画は今までに見たことがないくらい素晴らしい作品。特に主演の俳優の演技が秀逸。また他の出演者もそれぞれ主演になれるくらいの豪華な顔ぶれ。流石に巨額を投じた作品だけのことはあるな。...

「素晴らしい」「秀逸」「豪華」が判断の根拠

なぜか当たり前のことを言っている
ように聞こえる。なるほど感なし。

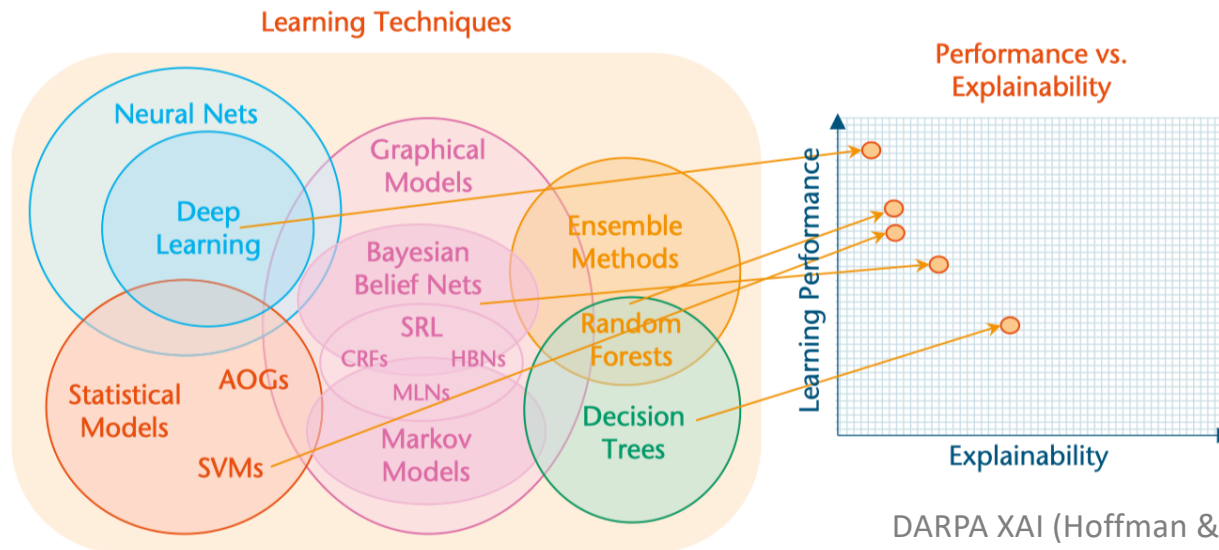
あなたの知ってるその知識とあの知識と
この新しい情報をかくかく組み合わせ
てこれこれの結果が得られたんですよ

みたいな説明？

人間の**知識に紐付けて説明**できる仕組みがほしい

深層学習でも未解決の広大な領域

- explainabilityとperformanceのトレードオフ
 - DARPA XAI のゴールはこのトレードオフを両立することだった
 - ただし、このトレードオフは、大量の訓練データが入手できて、end2endの深層学習がうまくいく領域での話



- しかし、世界にはこの条件が成り立たない課題が無数にある

「言語で説明できる」問題をねらいたい

- 人間が日常の中で**言語で説明している問題**の領域で、自分の判断を説明ができるXAIの設計論を研究したい
- 判断の過程や理由を言語で説明できるタイプの問題
 - 社会の至るところに存在する
 - 情報の真偽判断、法律的な判断、経営の意思決定など
 - 人がAIに説明の手本を示すことができる
 - 必要な知識が形式知としてどこかに記述されている

ゴール感

① 質問応答 / 文章読解

Context: *The princess climbed out the window of the high tower and climbed down the south wall when her mother was sleeping. She wandered out a good ways. Finally she **went into the forest** where there are no electric poles but where there are some caves.*

Question: *Where did the princess wander to after escaping?*

Answer: A) Mountain * B) Forest
C) Cave D) Castle

(MCTest, mc160.dev.29)

② 自然言語アセスメント

- 5 Johnny is testing water from a nearby creek to see how building a new golf course upstream is changing the water supply. Water is collected from the creek and placed in glass beakers. He then uses a hand lens to see what small animals live in the water. Afterwards, the water is tested for chemicals. Which safety procedure should be followed during the investigation?
- A** Clean up broken glass before telling the teacher.
- B** Mix tap water into the creek water so the sample is cleaner.
- C** Wear sandals so shoes do not get
- D** This is a safety procedure.
- E** Poor quality. Has bad science procedures. Water would be collected in sealable jars or bottles. Fourth grade students would not be using glass beakers for an outdoor investigation.
- F** By the fourth grade, students know that there will be microscopic organisms in the water. Possibly some macroscopic organisms that could be viewed with a hand lens, but if one is adding rigor, they would not be called small animals.
- G** The question indicates that safety procedure will be listed and the student is to pick the best one for this investigation. A, B, and C are not safety procedures.

The question indicates that safety procedure will be listed and the student is to pick the best one for this investigation. A, B, and C are not safety procedures.

D. This is a safety procedure.

「貨幣による意志の買収」という主題を明示する。①

「暴力的」を打ち消す抑揚的表現で説明する。③

貨幣は本質的に暴力の手段と、私たちが日常的に貨幣の影響を受け、自らの意志とは異なるか、あるいは、貨幣の持つ、ゆるぎのない教養化し、基準をという機能は、人間社会をなす諸原理の根本とせざるを得ない。

共通の尺度の具体例
行政の諸原理など

貨幣が、人間同士の共通の尺度になる点を押さえる。②

貨幣のどのような面が行動に影響を与えるのかと明確にする。④

傍聴部は直前の段落を受け、人間の意志が貨幣によって可能性を示したものの、その後には説明に着目する。

自らの意志によって、押し進められているように見えるが、影響を受けて、自然と自らの利益になる方向に移動する。⑤

「点を押さえる」

「なぜ何に束縛されているのか（影）」

「主観の原理に従って行動している」

「意識の無意識」という点を押さえる

言語で説明できるAIへの道のり

～ 質問応答、自然言語アセスメントを例題として ～

講演内容

1. イントロ: Explainable AI (XAI)
2. DNNと記号推論の融合を通して
3. マイルストーン: タスク設計とデータ構築

基本戦略

「知識に紐付けた」説明を

DNN/記号推論のハイブリッドで実現する

DNN推論
(システム1)

記号推論
(システム2)

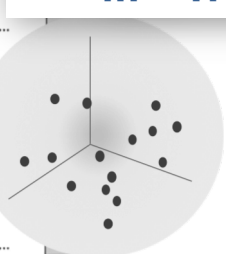
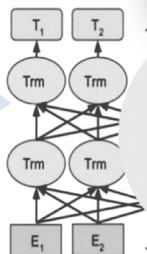
DNN推論と記号推論の
相互作用によって探索する

◎DNN推論の利点

全訓練データを
暗黙的に考慮
→ 推論が高速

◎DNN推論の利点

記号(言語表現)
の分散表現
→ 柔軟い推論



注意

教示

探索の
ボトルネック

知識獲得の
ボトルネック

知識

Text

KB

言語で書かれた知識を
DNNで柔軟かく推論に使う

◎記号推論の利点

推論の構造を
説明と見なせる

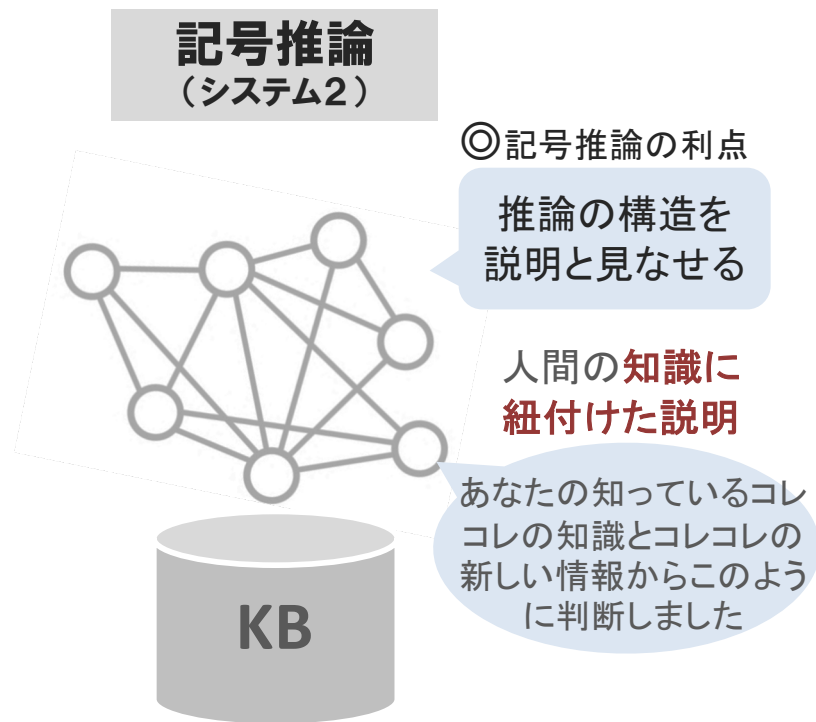
人間の**知識に
紐付けた説明**

あなたの知っているコレ
コレの知識とコレコレの
新しい情報からこのよう
に判断しました

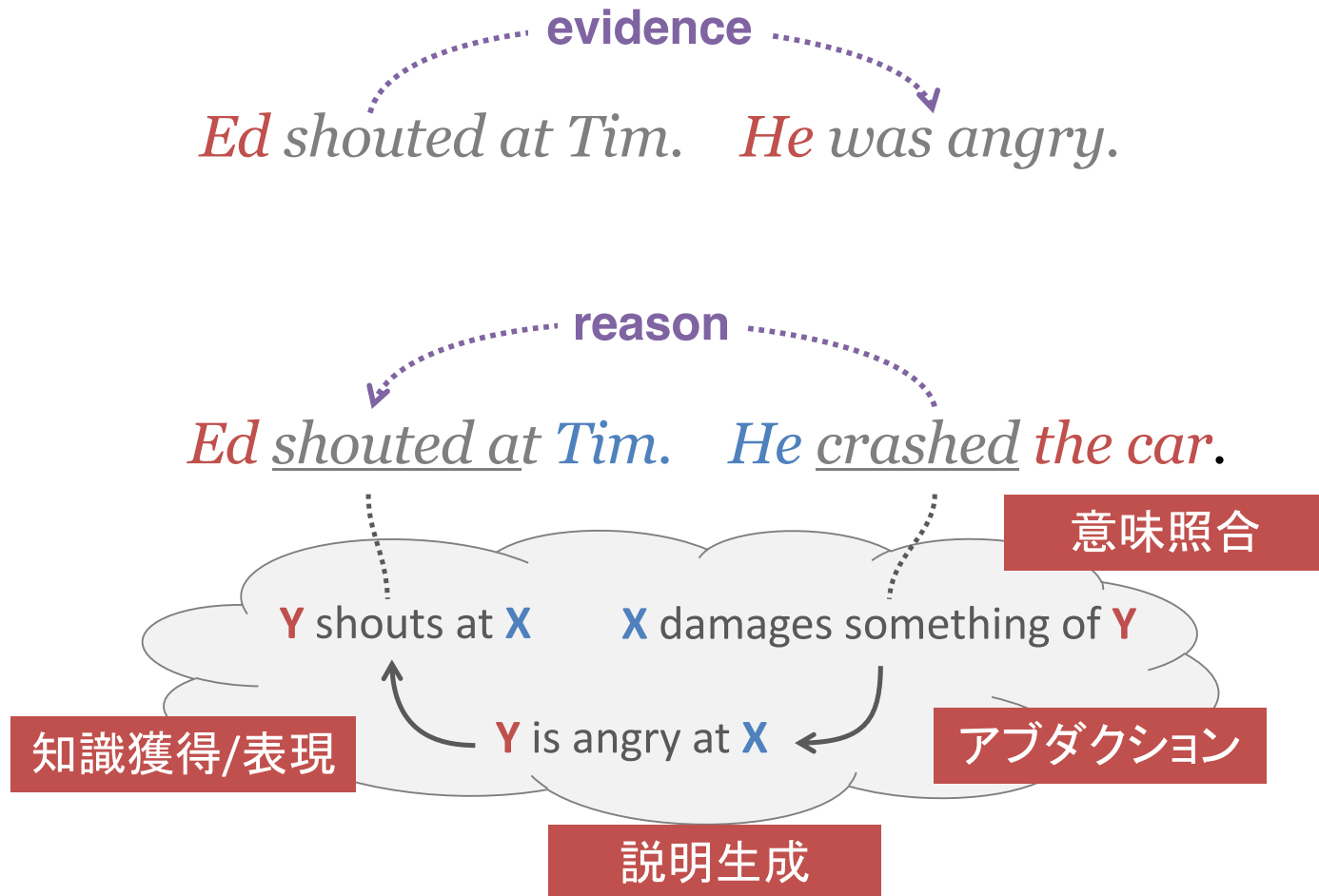
記号推論とDNN推論の相互作用をどう設計するかという新しい問い

基本戦略

「知識に紐付けた」説明を DNN/記号推論のハイブリッドで実現する



言語理解における知識と推論の問題



基本戦略

「知識に紐付けた」説明を DNN/記号推論のハイブリッドで実現する

記号推論 (システム2)

◎記号推論の利点

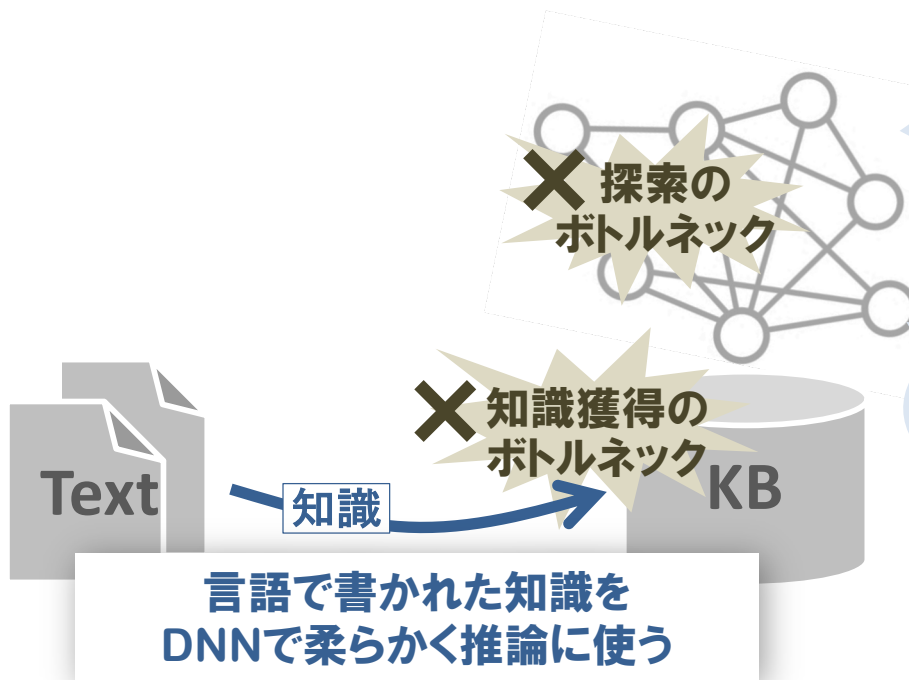
推論の構造を
説明と見なせる

人間の**知識に
紐付けた説明**

あなたの知っているコレ
コレの知識とコレコレの
新しい情報からこのよう
に判断しました

◎DNN推論の利点

記号(言語表現)
の分散表現
→ 柔軟かい推論

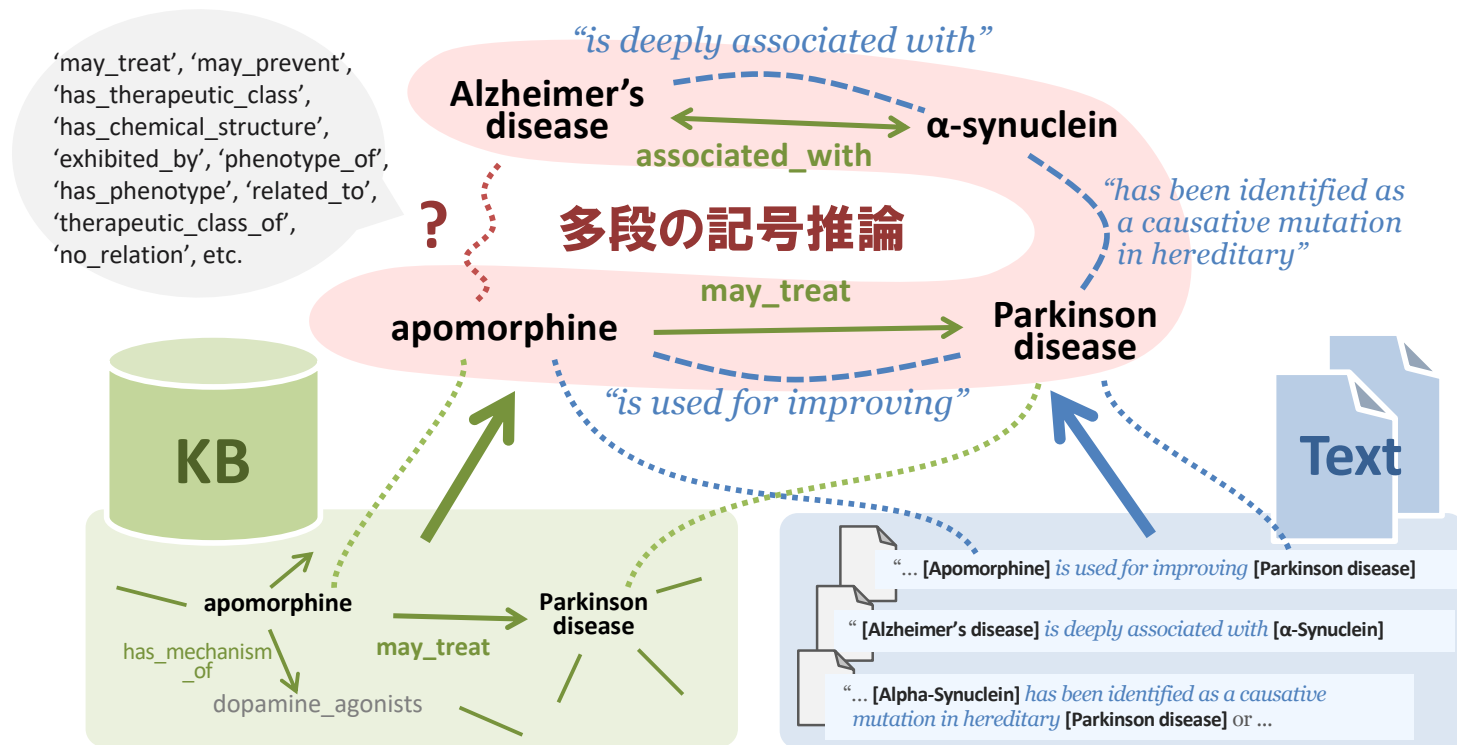


ひとつの例：Universal KG 上の推論

- 複数の関係を組み合わせる推論
- 言語で書かれた知識を併用
- 推論を記号(言語)で説明

人間が持っている知識に紐付けて判断を説明

アポモルフィン¹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²が関わっている(論文)と...

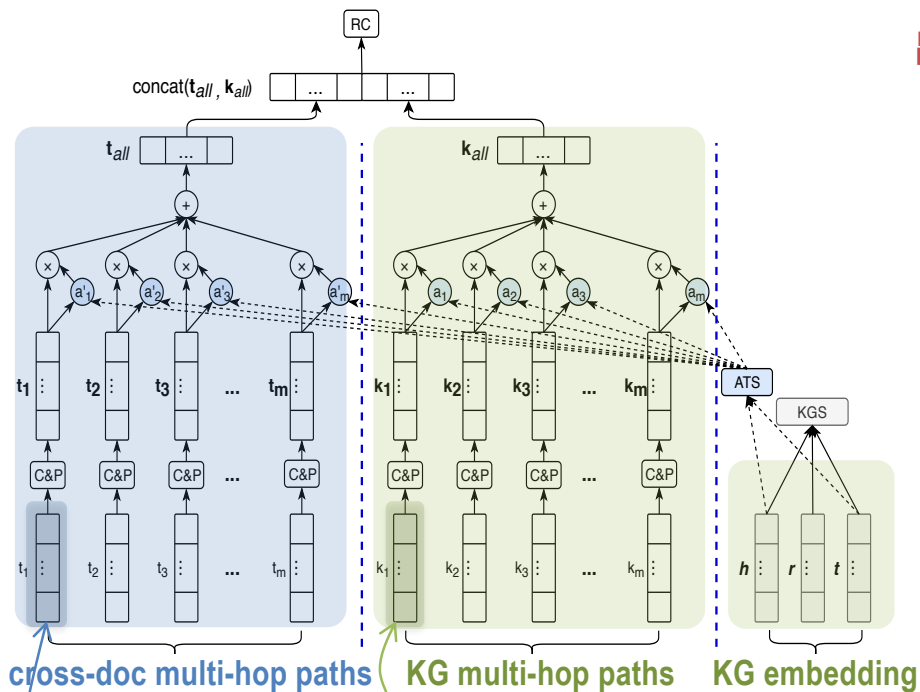




TOHOKU
UNIVERSITY

Relation prediction over UKG

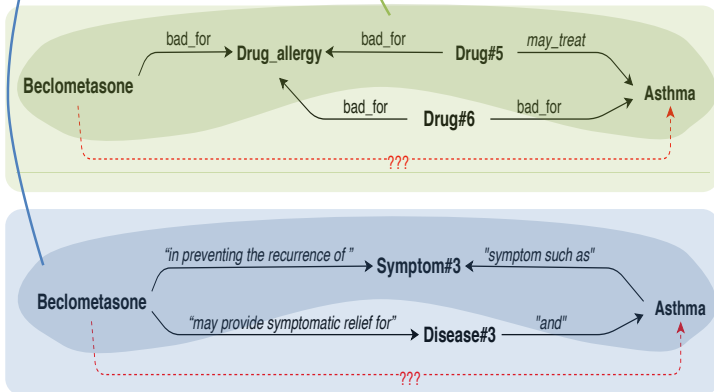
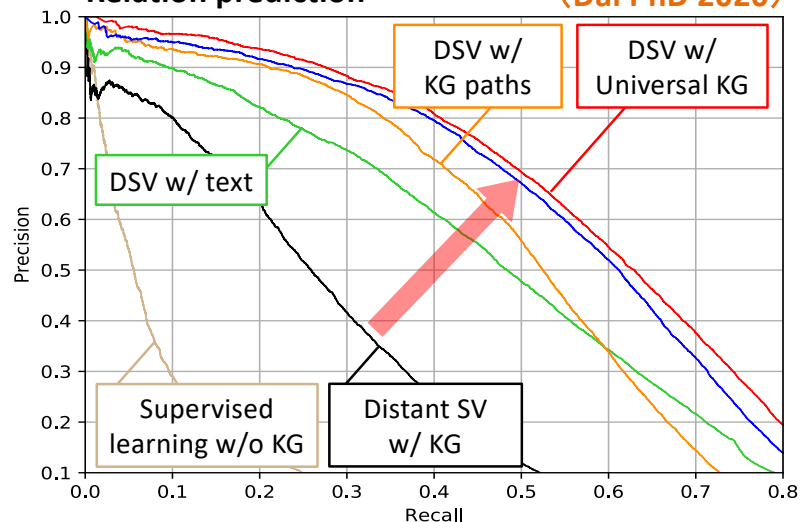
(Dai+ NLP 2020; PhD 2020; under review)



出力の説明性だけでなく知識抽出精度も大幅向上

Relation prediction

(Dai PhD 2020)



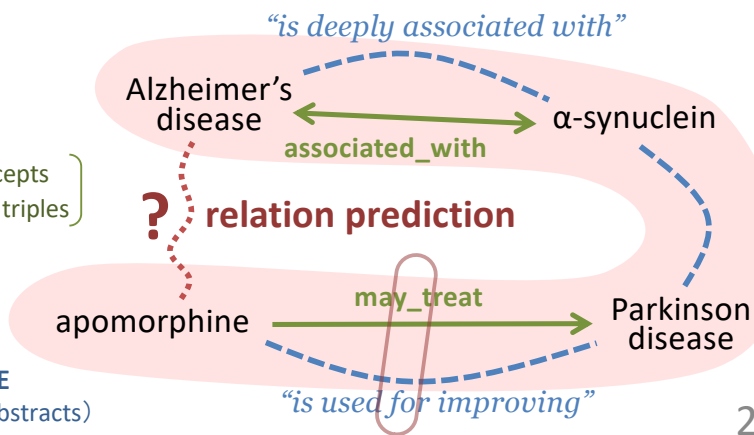
UMLS

[37K concepts
185K rel triples]



MEDLINE

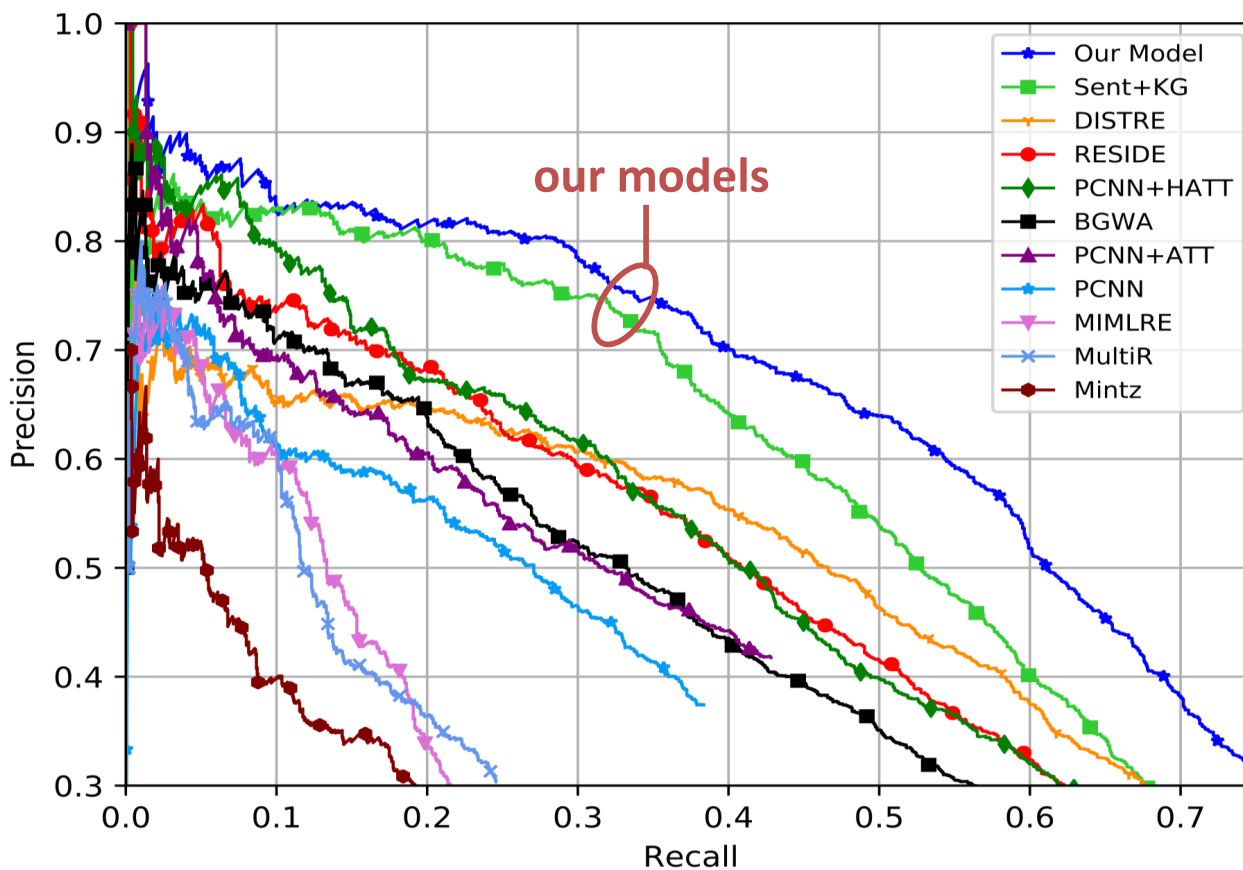
(24.3M abstracts)



Relation prediction over UKG

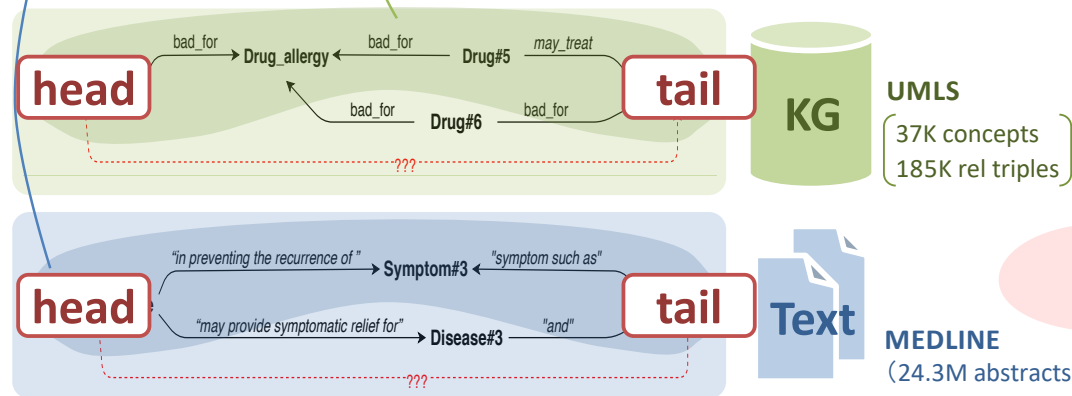
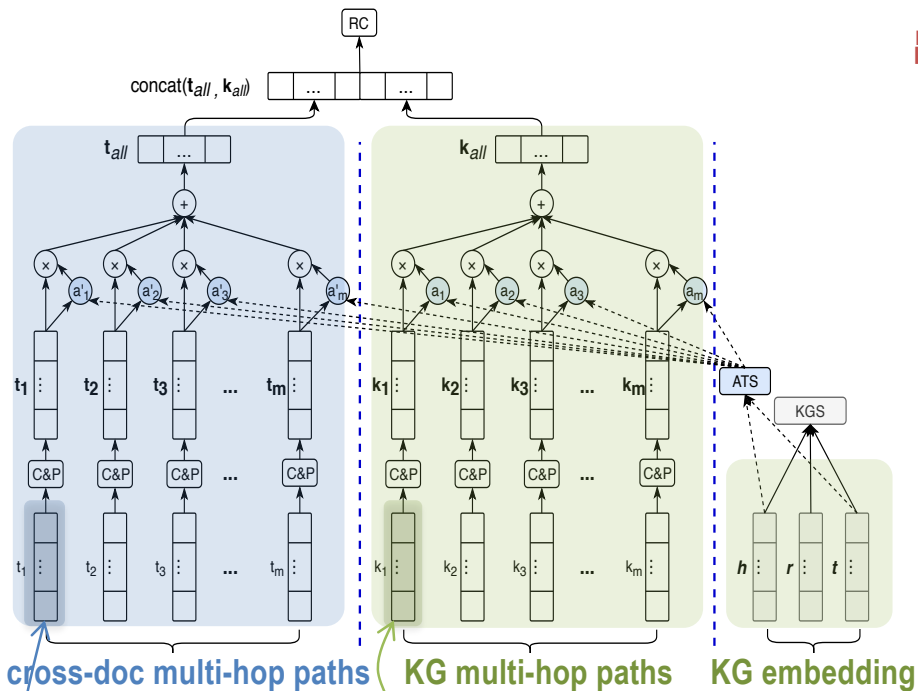
(Dai+ NLP 2020; PhD 2020; under review)

標準的なデータセット NYT10 でも SoTA

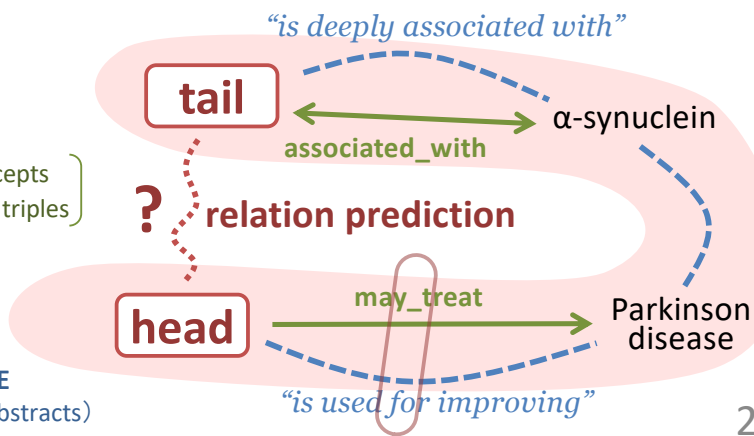
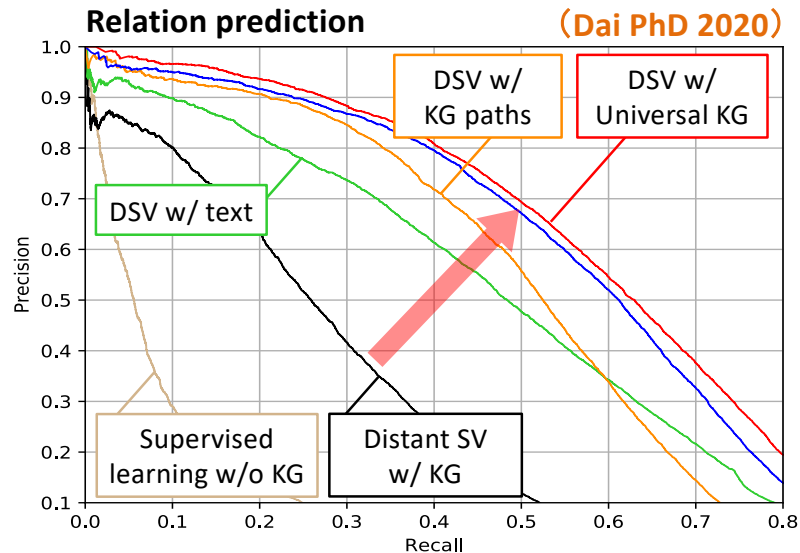


Relation prediction over UKG

(Dai+ NLP 2020; PhD 2020; under review)



出力の説明性だけでなく知識抽出精度も大幅向上



基本戦略

「知識に紐付けた」説明を

DNN/記号推論のハイブリッドで実現する

DNN推論
(システム1)

記号推論
(システム2)

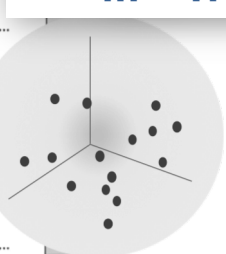
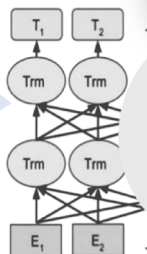
DNN推論と記号推論の
相互作用によって探索する

◎DNN推論の利点

全訓練データを
暗黙的に考慮
→ 推論が高速

◎DNN推論の利点

記号(言語表現)
の分散表現
→ 柔軟い推論



注意

教示

探索の
ボトルネック

知識獲得の
ボトルネック

知識

Text

KB

言語で書かれた知識を
DNNで柔軟かく推論に使う

◎記号推論の利点

推論の構造を
説明と見なせる

人間の知識に
紐付けた説明

あなたの知っているコレ
コレの知識とコレコレの
新しい情報からこのよう
に判断しました

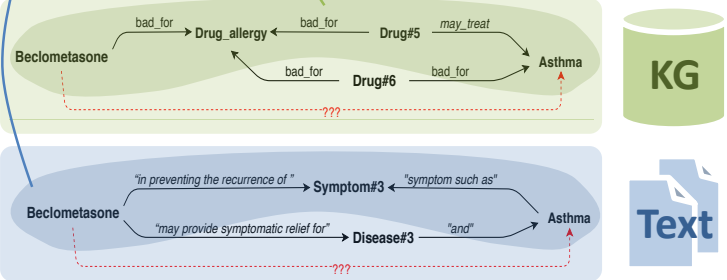
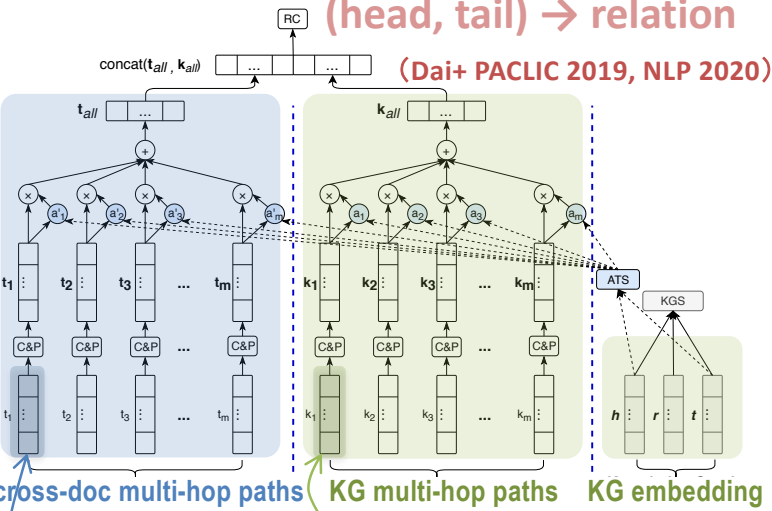
記号推論とDNN推論の相互作用をどう設計するかという新しい問い



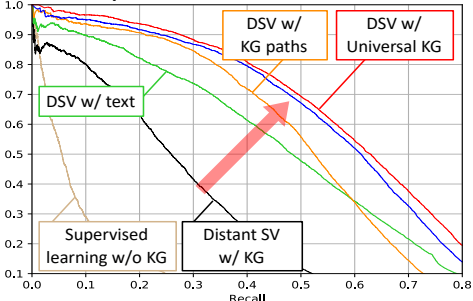
Dual Process Model

記号表現に基づく熟考的推論 (システム2)

(head, tail) $\rightarrow$ relation

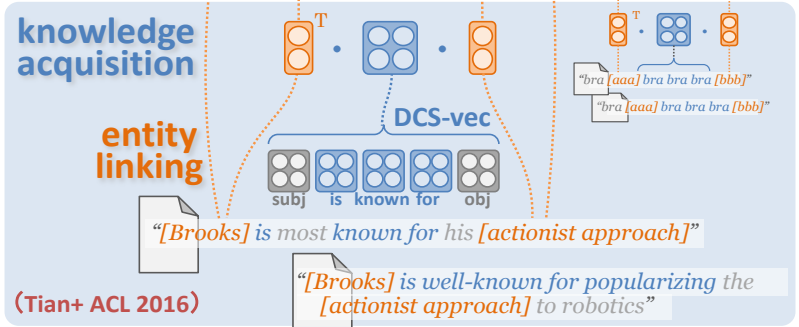
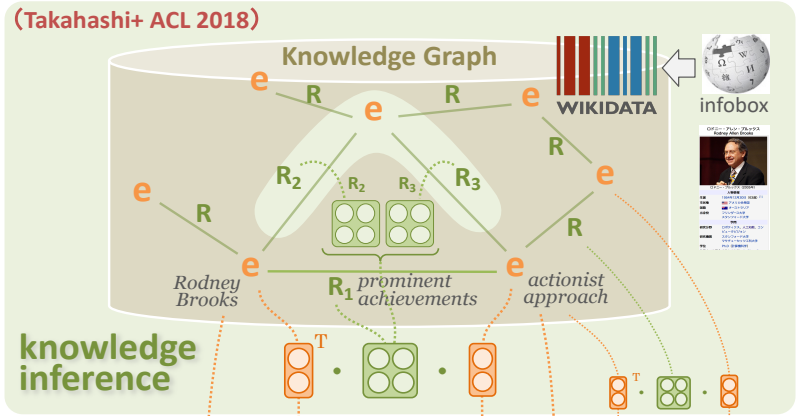


Relation prediction (Dai+ PACLIC2019, PhD2020)

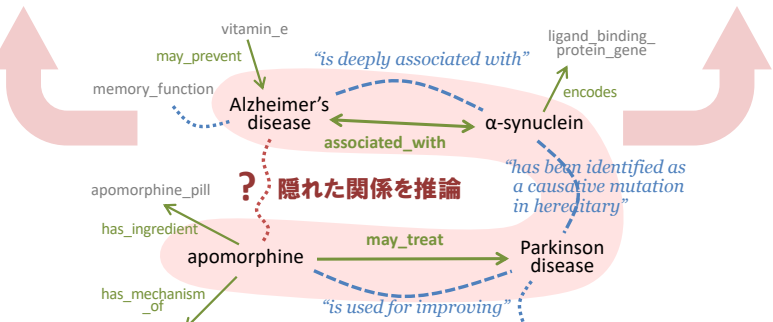


埋め込み空間における即応的推論 (システム1)

(head, relation) $\rightarrow$ tail



(高橋+ NLP2020, Heinzerling+ arXiv 2020)

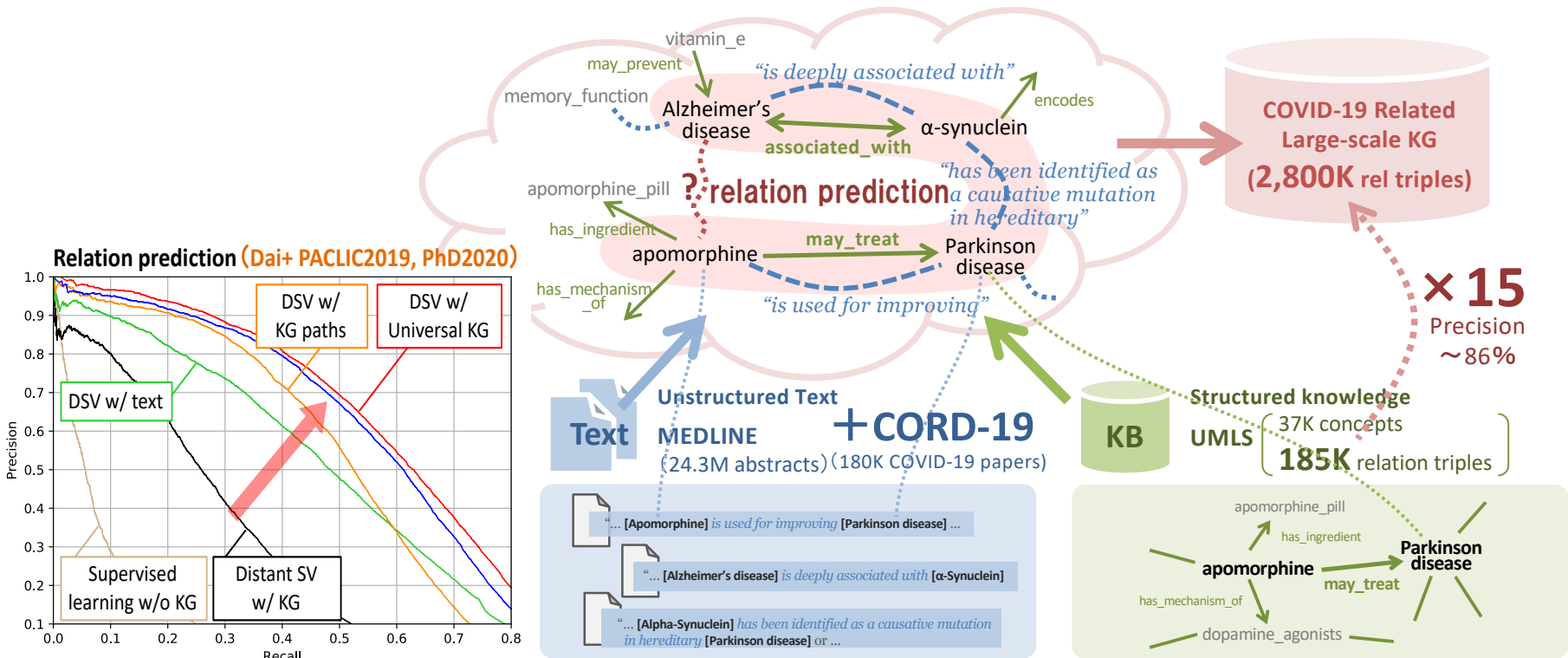


Building Large-scale Knowledge Graph from Biomedical Scientific Papers

Dai, Takahashi, Inoue, Inui (Tohoku U), Tsuruoka (U Tokyo), Aizawa (NII)

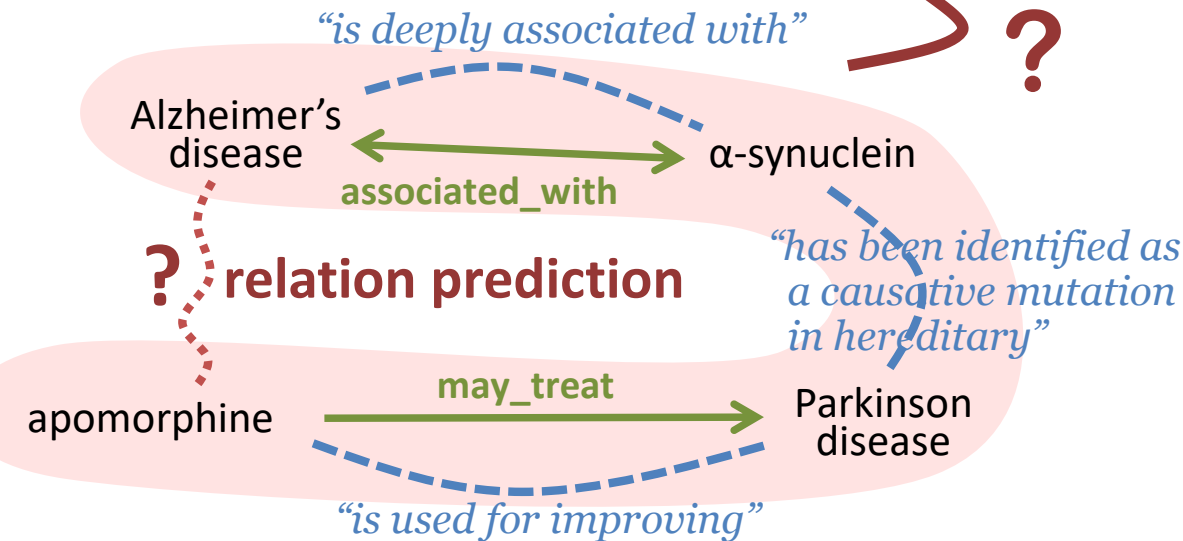
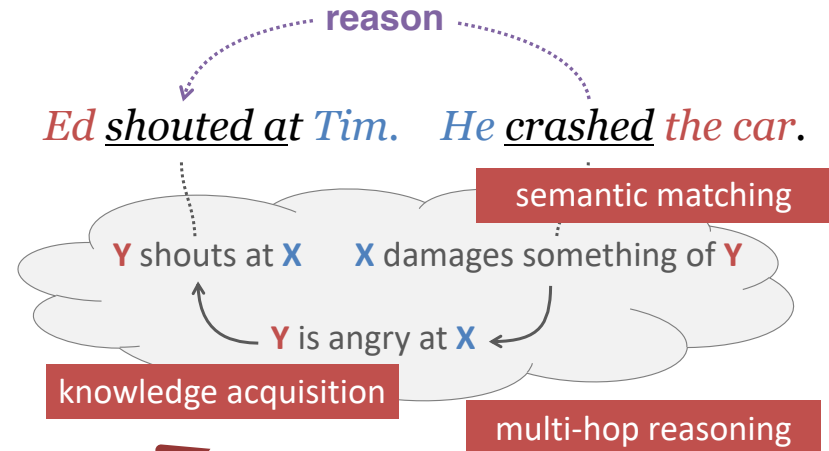
Multi-hop reasoning over biomedical scientific papers (text) combined with existing knowledge graph (KG):

- Extract stated (but unstored) relations from text
- Predict undescribed/undiscovered new relations



Next steps

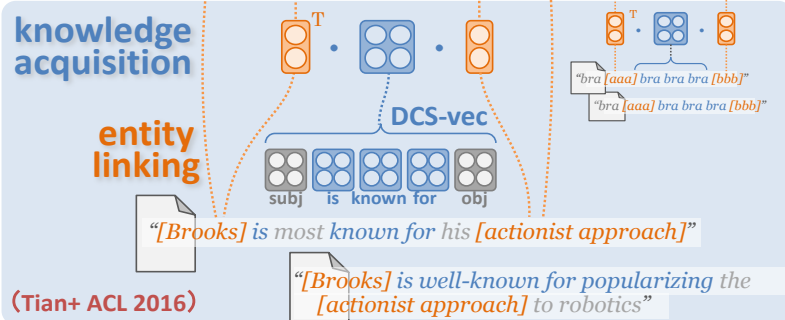
- エンティティ間の関係
→ イベント間関係
- 否定、条件、時間、等々



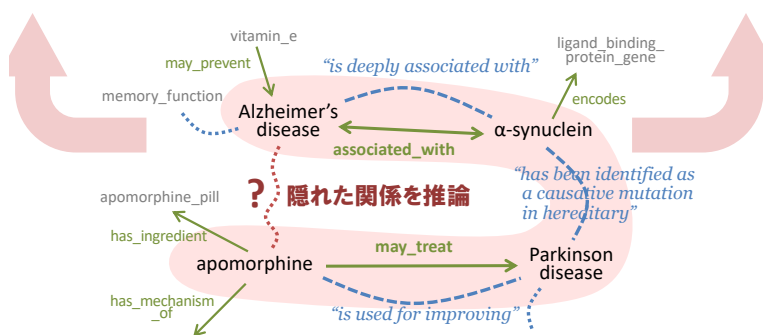


埋め込み空間における即応的推論（システム1）

(head, relation) → tail



(高橋+ NLP2020, Heinzerling+ arXiv 2020)



Language Models as Knowledge Bases: On Entity Representations, Storage Capacity, and Paraphrased Queries

(Heinzerling+ arXiv 2020)

Benjamin Heinzerling^{1, 2} and Kentaro Inui^{2, 1}

¹RIKEN AIP & ²Tohoku University

(Barack_Obama, wasBornIn, Hawaii)

Fact

Statement

Barack Obama was born in Hawaii.

DNN Language Model

Query

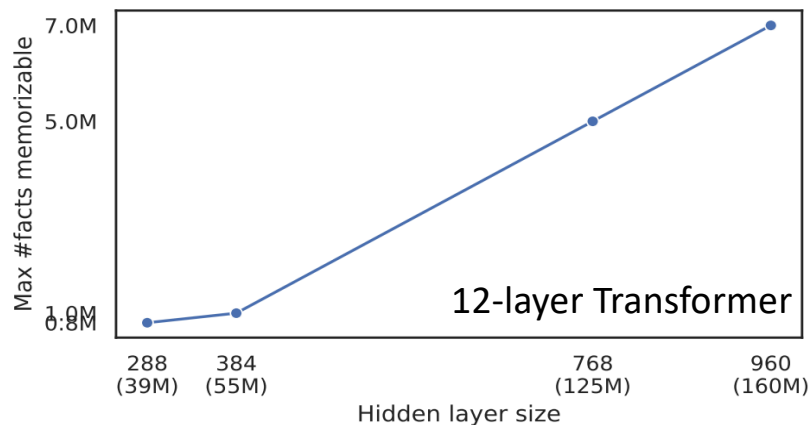
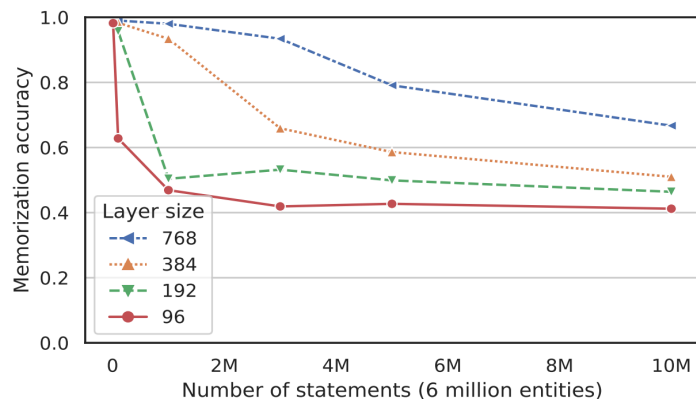
Barack Obama is from [MASK].

Answer

Barack Obama is from Hawaii.



Analysis of LM capacity for storing world knowledge facts:



言語で説明できるAIへの道のり

～ 質問応答、自然言語アセスメントを例題として ～

講演内容

1. イントロ: Explainable AI (XAI)
2. DNNと記号推論の融合を通して
3. マイルストーン: タスク設計とデータ構築

どんなマイルストーンを設定するか

① 質問応答 / 文章読解

Context: *The princess climbed out the window of the high tower and climbed down the south wall when her mother was sleeping. She wandered out a good ways. Finally she **went into the forest** where there are no electric poles but where there are some caves.*

Question: *Where did the princess wander to after escaping?*

Answer: A) Mountain * B) Forest
C) Cave D) Castle

(MCTest, mc160.dev.29)

② 自然言語アセスメント

5 Johnny is testing water from a nearby creek to see how building a new golf course upstream is changing the water supply. Water is collected from the creek and placed in glass beakers. He then uses a hand lens to see what small animals live in the water. Afterwards, the water is tested for chemicals. Which safety procedure should be followed during the investigation?

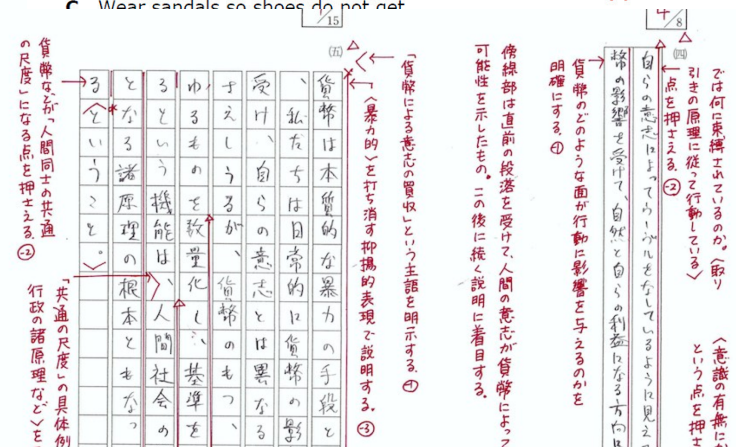
- A Clean up broken glass before telling the teacher.
- B Mix tap water into the creek water so the sample is cleaner.
- C Wear sandals so shoes do not get

Poor quality. Has bad science procedures. Water would be collected in sealable jars or bottles. Fourth grade students would not be using glass beakers for an outdoor Investigation.

By the fourth grade, students know that there will be microscopic organisms in the water. Possibly some macroscopic organisms that could be viewed with a hand lens, but if one is adding rigor, they would not be called small animals.

The question indicates that safety procedure will be listed and the student is to pick the best one for this investigation. A, B, and C are not safety procedures.

D. This is a safety procedure.



Handwritten analysis of the safety procedure question. The grid contains Japanese characters, and the surrounding text provides a detailed explanation of the reasoning for selecting option D as the correct safety procedure.

推論過程を問うQAタスク

$\mathcal{R}^4C$: Right for the Right Reasons RC

(Inoue+ NLP 2020, ACL 2020)

Input

Question

What was the former band of
the member of Mother Love Bone
who died just before the release of “Apple”?

Articles

Title: Return to Olympus

- [1] Return to Olympus is the only album by the alternative rock band Malfunkshun.
- [2] It was released after the band had broken up and after lead singer Andrew Wood (later of Mother Love Bone) had died...
- [3] Stone Gossard had compiled...

Title: Mother Love Bone

- [4] Mother Love Bone was an American rock band...
- [5] The band was active from...
- [6] Frontman Andrew Wood’s personality and compositions helped to catapult the group to...
- [7] Wood died only days before the scheduled release of the band’s debut album, “Apple”, thus ending the...

Output

Derivation (explanation)

[Malfunkshun]
is former of
[Mother Love Bone]

[Andrew Wood]
is a member of
[Mother Love Bone]

[Andrew Wood]
died just before the
release of [Apple]

- Set of entity relations
- Entities & relations are written in free-form NL

Answer

Malfunkshun

情報信憑性判断の支援

日々生産・拡散される誤情報

週刊 ネット上の情報検証まとめ

今週の“要注意”情報

- ✓ 「PCR検査受けた時点で団信入れず住宅ローン組めなくなる」
- ✓ 「厚労省HPからマスクのマークが消えた」
- ✓ 「中国の公人に日本の税金を支払う福岡市」
- ✓ 「菅官房長官公式アカウント」
- ✓ 「モネは40代で画業に入った」

Vol.51/ 2020.9.23

福岡市 反日行動になりますよ。
税金の無駄をはるかに超えて、
日本侵略を公言するチャイナ共産党の工作員である公
人に日本の税金を支払うとは、
福岡市終わってます。市民には反対してほしい。— 華
字メディア recordchina.co.jp/newsinfo.php?i...
@recordchinaより



福岡市が中国の公務員800人規模の受け入れを発表、名目上は「研修」だが、一
4日、中国では昨今、公務員の「海外研修」が盛んだが、その実態は「研修名目
の旅行」というのは公然の秘密。そんな折、福岡市が来年4月から中国の公務員...
© recordchina.co.jp

信憑性判断（ファクトチェック）を説明可能AIで支援

ツイート

現在、1万人近くの感染者がいるが、日本人は4226人。半分もない。
感染者の多くが外国人で、日本人が治療を受けられない。#治療は日本人を優先

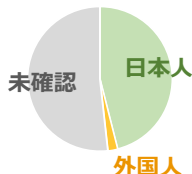
信頼できるか？
その根拠は？

厚生労働省Web

国内感染者は9027人。
うち日本国籍が確認されている者4226人。
外国籍が確認されている者52人。

メディア

- 感染者の約半数は国籍未確認。
- 厚労省は「未確認の方のほとんどが日本人と推測される」と述べた。



割合の計算

$$1.0 - \frac{4226 + 52}{9027} \approx 0.5 \Rightarrow \text{約半数が国籍未確認}$$

知識の参照

「日本国籍」と「外国籍」の相補的関係

ユーザの
内部状態を予測

未確認の数？

日本人は半分以上？以下？
信頼できそう？



ユーザ

このツイート、信じていい？

..... という厚労省の情報と整合し
ません。そう判断した理由は

でも残りはみんな外国人の可
能性もあるのでは？

いえ、厚労省の推測では

信頼できる(公的)情報源を探索

言語理解・推論

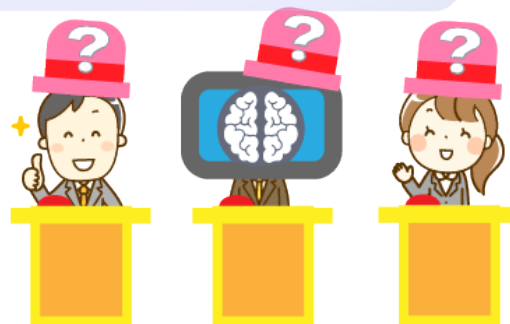
信憑性
判断支援AI

知識と推論に基づく説明可能AI × 説明のコミュニケーション

クイズAIコンペを主催

Q: 北海道の中心に位置することから「北海道のへそ」を名乗る、ラベンダーで有名な都市はどこ？

AI王 クイズ AI 日本一決定戦



現在の結果 [重要日程・ルール](#) [投稿ファイルの仕様](#) [新規投稿](#)

JAQKET

#	Model	Code	Accuracy	(easy / hard / very hard)
1 02 Nov 2020	kat IR(TF-IDF)x2, BertForMultipleChoice, Heuristic, NoDev1/2	N/A	0.948	(0.98 / 0.68 / 0.30)
2 29 Sep 2020	二宮チーム (NAISTデータ駆動知識処理研究室/NICT) IR(TF-IDF), NICT BERT + additional pretraining (wikipedia), model ensemble	N/A	0.936	(0.97 / 0.48 / 0.15)
3 06 Nov 2020	Takahiro Aso test	N/A	0.929	(0.96 / 0.55 / 0.22)
4 20 Nov 2020				0.45 / 0.04)
5 09 Jun 2020				0.35 / 0.22)
6 01 Nov 2020	BERT	N/A	0.902	(0.94 / 0.39 / 0.07)

2020年11月現在のランキング
(リーダーボード)

BERTに基づくベースラインを配布中
どなたでも簡単にご参加いただけます

2021年3月に最終順位決定戦を開催予定
まだ4ヶ月、これからでも間に合います！

くわしくは:

AI王



どんなマイルストーンを設定するか

① 質問応答 / 文章読解

Context: *The princess climbed out the window of the high tower and climbed down the south wall when her mother was sleeping. She wandered out a good ways. Finally she **went into the forest** where there are no electric poles but where there are some caves.*

Question: *Where did the princess wander to after escaping?*

Answer: A) Mountain * B) Forest
C) Cave D) Castle

(MCTest, mc160.dev.29)

② 自然言語アセスメント

5 Johnny is testing water from a nearby creek to see how building a new golf course upstream is changing the water supply. Water is collected from the creek and placed in glass beakers. He then uses a hand lens to see what small animals live in the water. Afterwards, the water is tested for chemicals. Which safety procedure should be followed during the investigation?

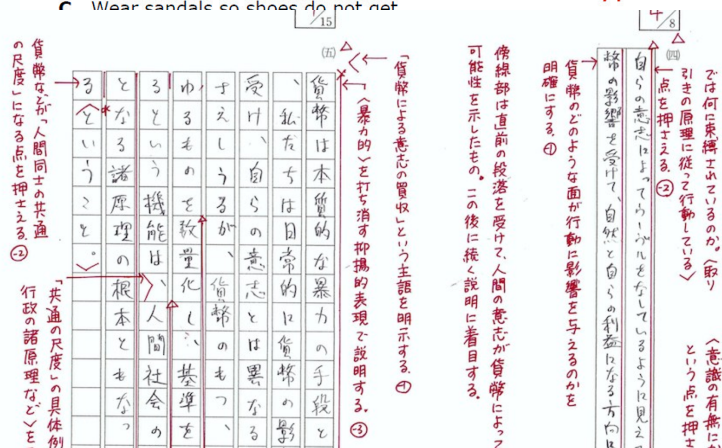
- A Clean up broken glass before telling the teacher.
- B Mix tap water into the creek water so the sample is cleaner.
- C Wear sandals so shoes do not get

Poor quality. Has bad science procedures. Water would be collected in sealable jars or bottles. Fourth grade students would not be using glass beakers for an outdoor Investigation.

By the fourth grade, students know that there will be microscopic organisms in the water. Possibly some macroscopic organisms that could be viewed with a hand lens, but if one is adding rigor, they would not be called small animals.

The question indicates that safety procedure will be listed and the student is to pick the best one for this investigation. A, B, and C are not safety procedures.

D. This is a safety procedure.



Handwritten analysis of the safety procedure question. The grid contains Japanese characters, and the surrounding text provides a detailed explanation of the reasoning for selecting option D as the correct safety procedure.

AIP 理研AIPセンター 自然言語理解チーム

テーマ：自然言語アセスメント

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



例① 記述式答案の採点・解説 (短中期課題)

① 次の文章を読んで、後の問いに答えよ。
西洋文化の基底の「対決」は神対人間(宗教=契約)、人間対自然(科学=合理主義)、人間対人間(個人主義)という形で現れるということ。日本人のように余計なことを言わないのではなく、他人に分かってもらうために言葉を尽くす「対決」のスタンスが西洋の文化を導いてきた。

問 傍線部(1)「こうした～築いてきた」とあるが、それはどういうことか。70字以内で説明せよ。

西洋文化の基底の「対決」は神対人間(宗教=契約)、人間対自然(科学=合理主義)、人間対人間(個人主義)という形で現れるということ。日本人のように余計なことを言わないのではなく、他人に分かってもらうために言葉を尽くす「対決」のスタンスが西洋の文化を導いてきた。

模範回答

減点対象

誤字 脱字 1点減点
文末が「こと」終わってないもの 1点減点

① 言葉を尽くして 3点
② 他人を説得する 3点

③ ①他人は自分と違うという説明 3点
② 対決のスタンス 2点
③ ②自分と異なる考え方を持つ 2点

④ ①他人は自分と違うという説明 3点
② 対決のスタンス 2点
③ ②自分と異なる考え方を持つ 2点

⑤ ①他人は自分と違うという説明 3点
② 対決のスタンス 2点
③ ②自分と異なる考え方を持つ 2点

⑥ ①他人は自分と違うという説明 3点
② 対決のスタンス 2点
③ ②自分と異なる考え方を持つ 2点

⑦ ①他人は自分と違うという説明 3点
② 対決のスタンス 2点
③ ②自分と異なる考え方を持つ 2点

⑧ ①他人は自分と違うという説明 3点
② 対決のスタンス 2点
③ ②自分と異なる考え方を持つ 2点

⑨ ①他人は自分と違うという説明 3点
② 対決のスタンス 2点
③ ②自分と異なる考え方を持つ 2点

⑩ ①他人は自分と違うという説明 3点
② 対決のスタンス 2点
③ ②自分と異なる考え方を持つ 2点

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.
Becouse my mother's hobby is gardening. And I beyond gardening club for 11 years ...*

《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".

<過去完了>は、ある過去の時点を示した上で、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.

《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)

Smoking **should be banned**! It causes **bad breath** and makes your **clothes stinky**!

While your point is **logically valid**, it **does not justify** banning the substance. Can you provide a more relevant reason?

Oh..I know that **many people die of second-hand smoke each day!**

Indeed, **people dying from second-hand smoke** is much more convincing than **causing bad breath**. Where did you learn about this information?





TOHOKU
UNIVERSITY



記述答案採点（国語）

（Mizumoto+ BEA 2019; Funayama+ ACL-SRW 2020; Ou+ under review）

日本語記述式問題自動採点デモ

次の文章を読んで、後の問いに答えよ。

西洋文芸の本質は（読舌）にある。それに対して日本文芸の本質は黙説（寡黙）にある。黙説とは聞かない言葉だが、省略、含意、暗示、ほのめかしなどを能うかぎり言葉を抑制した表現全体を指す。西洋人は基本的には他人は自分と異なる人間（異人）と見なす。だから自分の考えに他人を同意させる必要があると考える。他人に分かってもらうために言葉を尽くして自分の考えを伝えようとする。説得である。それに対して日本人は他人は自分とおなじ人間（同類）と考える。だから余計なことは言わない。暗黙の前提に寄りかかる。私にいわせれば、西洋文化の基底には「対決」のスタンスがある。その「対決」は神対人間（宗教＝契約）。人間対自然（科学＝合理主義）、人間対人間（個人主義）という形で現れる。こうした緊張のスタンスこそが読舌な西洋文化を導いてきたのである。（野内良三「偶然を生きる思想」による）

問: 下線部「こうした緊張のスタンスこそが読舌な西洋文化を導いてきた」とあるが、それはどういうことか。句読点とも七十字以内で説明せよ。

解答入力欄:

西洋人は基本的に他人は自分とは異なる人間と見なすので、言葉を尽くして自分の考えを伝え、他人に自分の考えを同意させる必要があること。

I

0文字

Submit

合計点:

12 / 16

西洋人は基本的に他人は自分とは異なる人間と見なすので、言葉を尽くして自分の考えを伝え、他人に自分の考えを同意させる必要があること。

採点基準A 基準 根拠事例

2 / 2

「西洋（では）」（＝話題の中心）・・・2点
※「ヨーロッパ」も可
※「外国」は不可

採点基準B 基準 根拠事例

3 / 5

「他人を自分とは異なる考え方をもつ人間と（みなす）」・・・5点
① 他人は自分とは違うという説明・・・3点
※「対決のスタンス」はBのポイント2点
② （異質な）考え方（を持つ）・・・2点
※B②の違いの中身の具体化

採点基準C 基準 根拠事例

3 / 3

自分の意見に同意を得るために・・・3点
※「他人にわかってもらうため」なども可

採点基準D 基準 根拠事例

4 / 6

「言葉を尽くして他人を説得する」・・・6点
① 「言葉を尽くして」・・・3点
※「手を替え品を替え」や「あれこれ工夫して」などは2点
② 「読舌」そのままは1点
③ 「（他人を）説得する」・・・3点
※「（自分の考えを）説明する」や「自分の考えを伝えようとする」は1点

誤字・脱字減点 基準

0 / 1

※ 誤字・脱字・転記ミス・主述のねじれなどは1点減点

文末減点 基準

0 / 1

※ 文末が「こと」、「事」でないものは1点減点

記述答案採点（英語）

フィードバック付き英作文自動採点デモ

次の日本語の下線部を英語に訳しなさい。ただし、解答に与えられた語句で文を始めること。

私が最初に英語を教わった先生はアイルランド人でした。彼女は日本で30年暮らしていて、日本語がペラペラでした。

解答入力欄:

She spoke Japanese well as she had lived in Japan for thurty years.

-語

Submit

She spoke Japanese well as she had lived in Japan for thurty years.

採点基準A

基準

アドバイス



「流暢な日本語」を適切に表現できる副詞として *fluently* などがあります。

採点基準B

基準

アドバイス



理由の従属接続詞 *because* を使いましょう。

採点基準C

基準

アドバイス



従属節内の時制が適切に書けている

採点基準D

基準

アドバイス



場所を表す表現(前置詞+名詞)が適切に書けている

採点基準E

基準

アドバイス



期間を表す語が適切に書けている

スペルミス

アドバイス



thurty → *thirty*.

論述文の論理解析とフィードバック生成

(Reisert+ AIED 2019; Naito+ 2020)

ディベートにおける反論の論理構造を内容レベルで解析し、フィードバックを生成

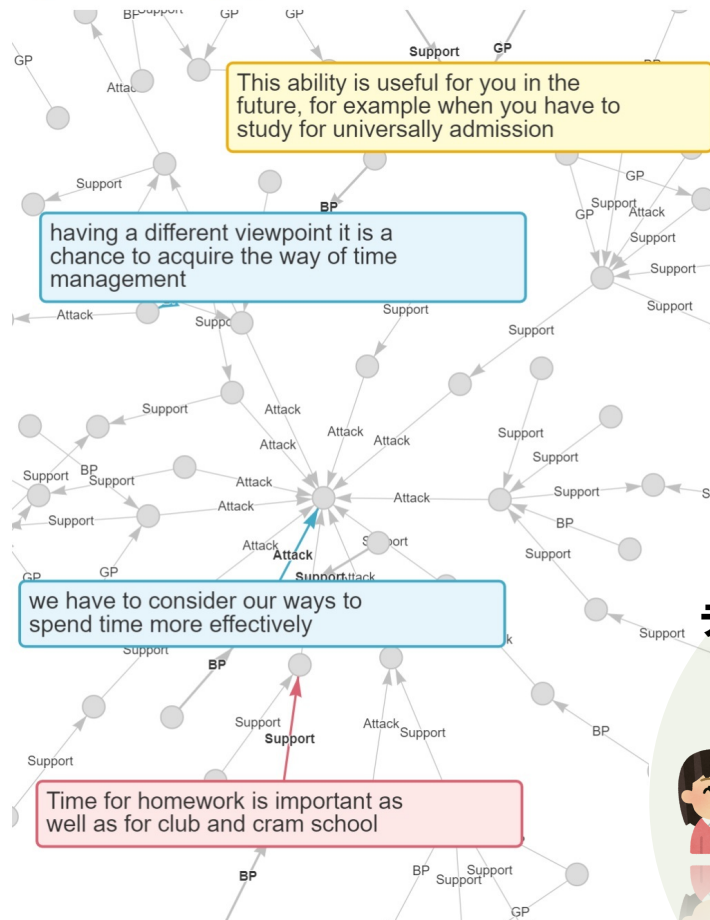
Speech of prime minister

Today's topic is "Homework should be abolished". We believe if homework were to be abolished, we could have more free time. As a result, we could do more of what we really wanted like club activities, hobbies, or playing with friends. In my case, I go to tennis club after class and then I go to cram school. After this full day, I start my homework at nearly 10:00 pm. I have a lot of homework, so I go to bed late at night and I don't have the opportunity to sleep for a long period of time. It is not healthy. Therefore, homework should be abolished.

Counter argument

I think club activities, hobbies, and playing with friends are necessary, but I also think that being able to plan ahead and be able to do homework is also a necessary ability. Rather than waiting for you to have some free time, you can improve your ability to manage time later in life by creating it yourself. So I think homework which I can manage and do should not be abolished.

Argumentation graph



Feedback

Suggestion:

タイムマネジメント力が将来にもいかに重要になってくるかで説明できるとさらによいでしょう

Suggestion:

どのように効率的に時間を使うか、具体的に話せるとよりよいでしょう。

Good Point:

時間が限られていることは必ずしも悪いことではないと反論できたのがよかったです

ディベートの対話的指導へ (中長期課題)

Smoking **should be banned**! It causes **bad breath** and makes your **clothes stinky**!

While your point is **logically valid**, it does **not justify** banning the substance. Can you provide a more relevant reason?

Oh...I know that **many people die of second-hand smoke** each day!

Indeed, **people dying from second-hand smoke** is much more convincing than **causing bad breath**. Where did you learn about this information?

Submit

Interactive tutoring

(Reisert+ AIED 2019; Naito+ 2020)

Constructive Feedback Demo

Topic

School Uniforms: Good or Bad?

I think uniforms are good. Uniforms can remove the wall of status. They oppose the spirit of unity and its celebration. They also look very nice and respectful.

Suggestion: Change 'oppose' to 'can oppose'

Suggestion: Add the phrase 'Based on my personal experience' to your argument.

Your argument: They also look very nice and respectful.

If you are making an argument based on your own experience, consider explicating this.

ねらい

・なぜ自然言語アセスメントか？

- **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**：①**説明の提供**、②**不全文の頑健解析**、③「**意図**」の推測
→ 自然言語処理の**未解決重要課題**（骨のある課題）

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：

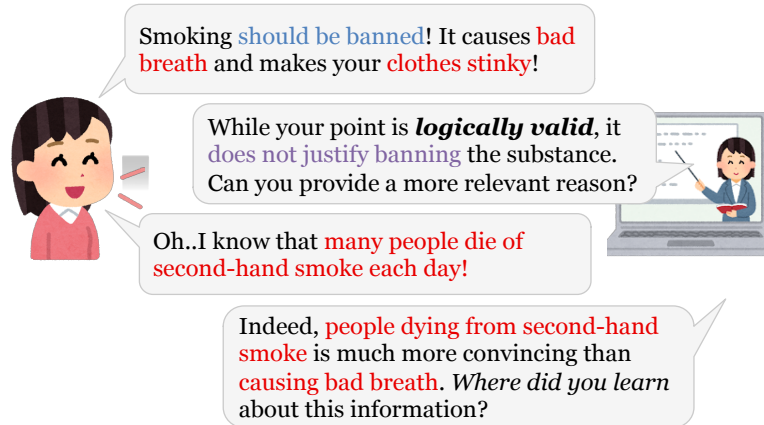
*I **had ever been to** gardening.
Becouse my mother's hobby is
gardening. And I **beyond**
gardening club for 11 years ...*

「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".

<過去完了>は、ある過去の時点を示した上で、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.

「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」と表現してみましょう。
<<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）



Smoking **should be banned**! It causes **bad breath** and makes your **clothes stinky**!

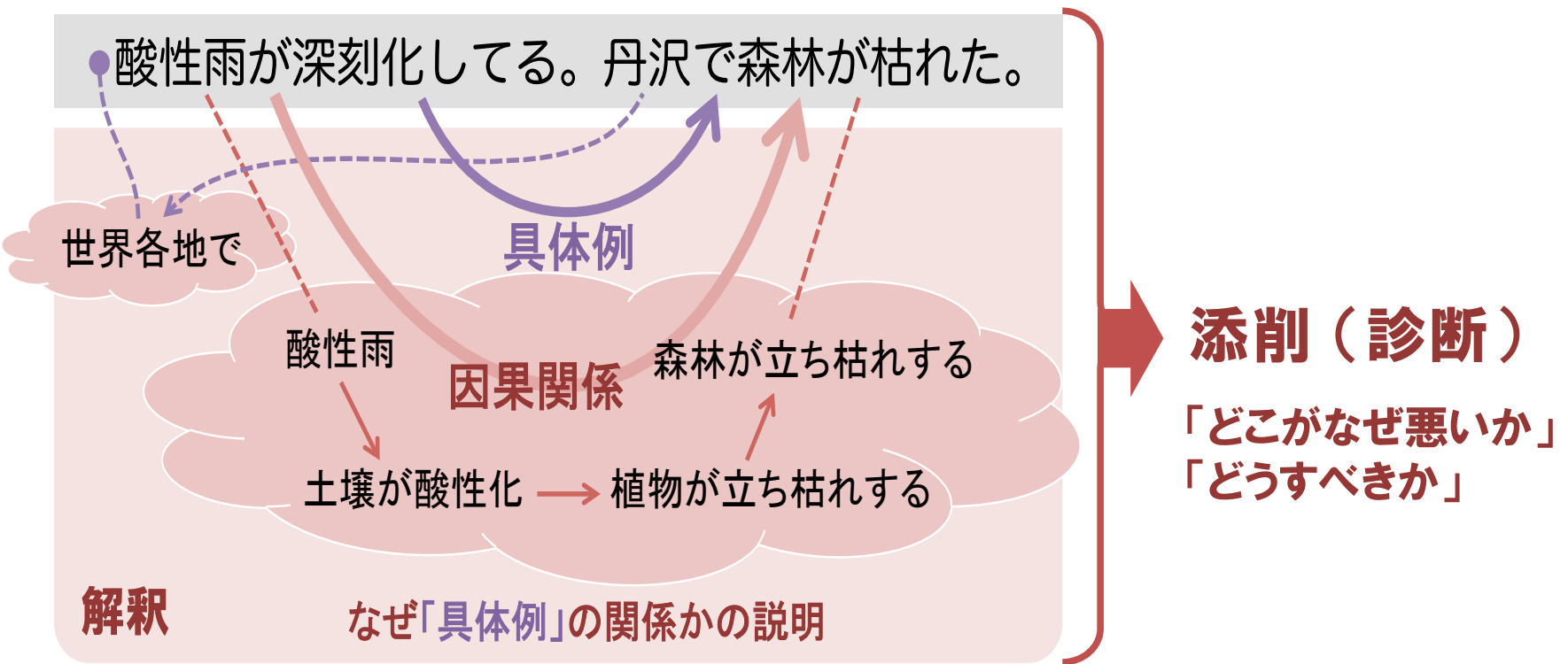
While your point is **logically valid**, it **does not justify banning** the substance.
Can you provide a more relevant reason?

Oh..I know that **many people die of second-hand smoke each day!**

Indeed, **people dying from second-hand smoke** is much more convincing than **causing bad breath**. *Where did you learn about this information?*

表現・内容の質の評価

「うまく表出できていない意味」を含め、文章の「言わんとする内容」を**解釈**して初めて、**添削(診断)**が可能になる



「行間を読む」言語理解

言明どうしを**知識**で繋ぎ合わせて文を超えた
繋がりを理解し、説明を含めた**解釈**を作る

ねらい

・なぜ自然言語アセスメントか？

- 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：①説明の提供、②不全文の頑健解析、③「意図」の推測
→ 自然言語処理の未解決重要課題(骨のある課題)

・ところが、研究利用可能なデータが決定的に不足

→ 研究者の参入少 = **ブルーオーシャン**

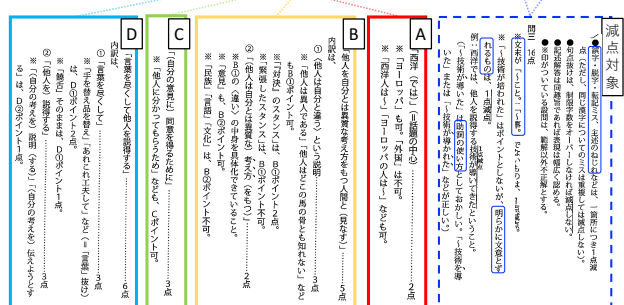
・データ収集問題を押さえれば、世界をリードできる可能性も

西洋文化の基底には様々な「対決」のスタンスがあり、そのスタンスとは神対人間、人間対人間のような緊張をあたえるもので、そのスタンスは「対決」のスタンスがある。その「対決」は神対人間、人間対自然、人間対人間という形で現れること

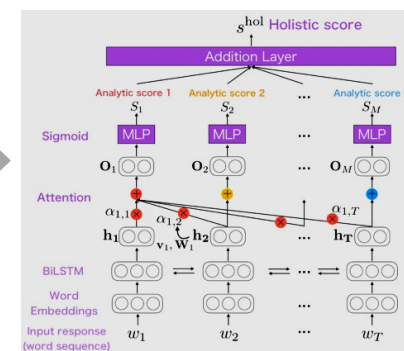
ユノロツは民族が言語や文化がせまく難解しており、何を考えているかも分からない異人に自分の考えを説明して認めさせないといけない

西洋文化の基底には「対決」のスタンスがあり、その「対決」は神対人間、人間対自然、人間対人間という形で現れること。

西洋の文化の基底には対決というスタンスがある。これが自分と他人は異人という考えを生み出し、説得になった。




技術の研究開発



多様な教育事業者と2階建て連携

理研AIPがハブとなってデータを共有化

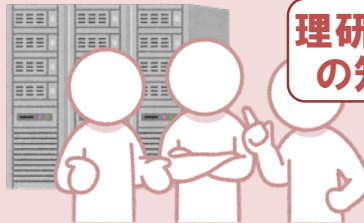
教育事業者から
データを収集

得られた技術は
理研単独の知財
として広く提供

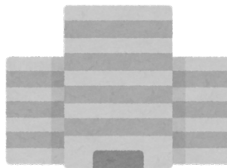
1 階



実データ提供
(問題、答案、基準、採点等)



**アセスメント技術
の研究開発**

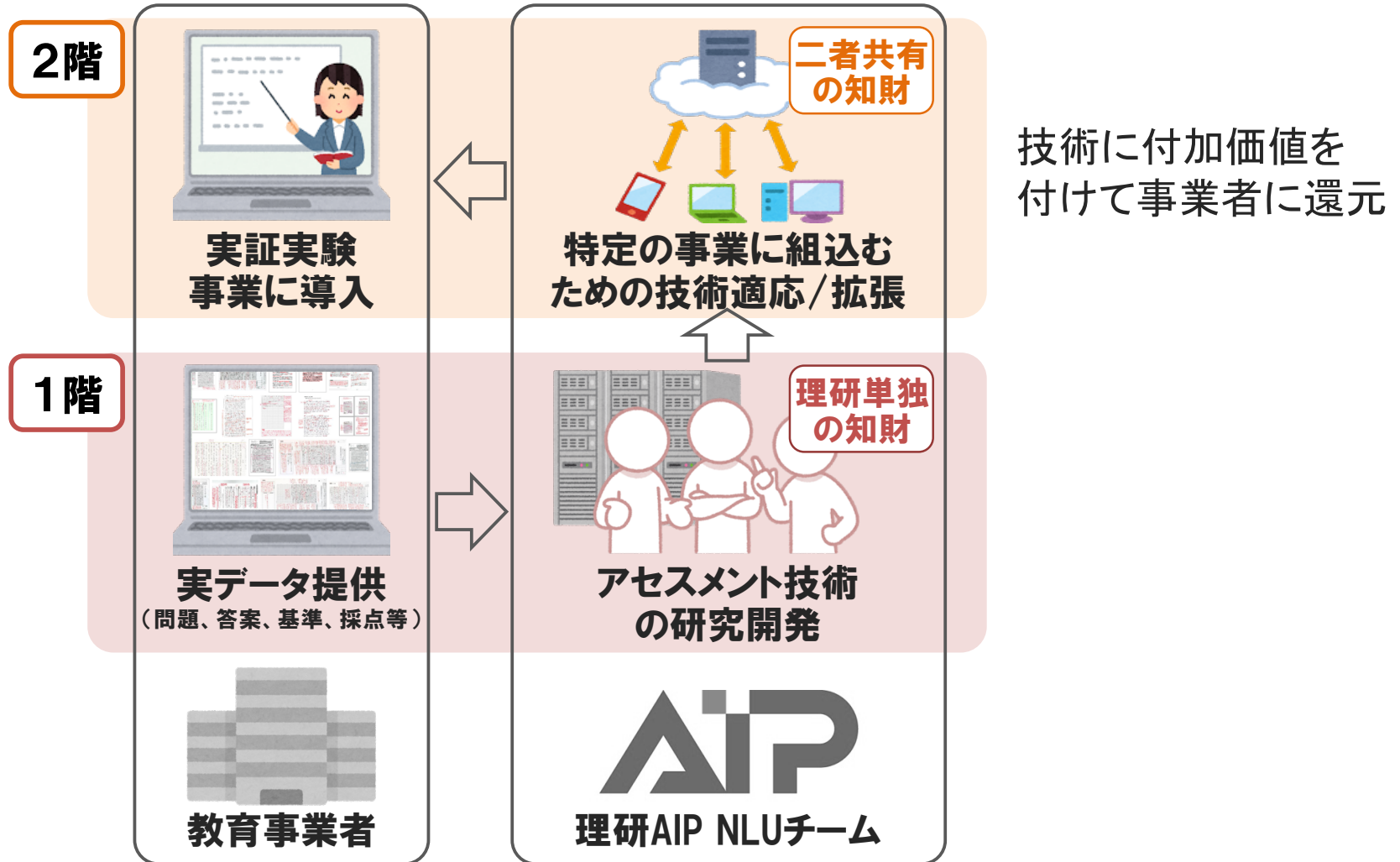


教育事業者

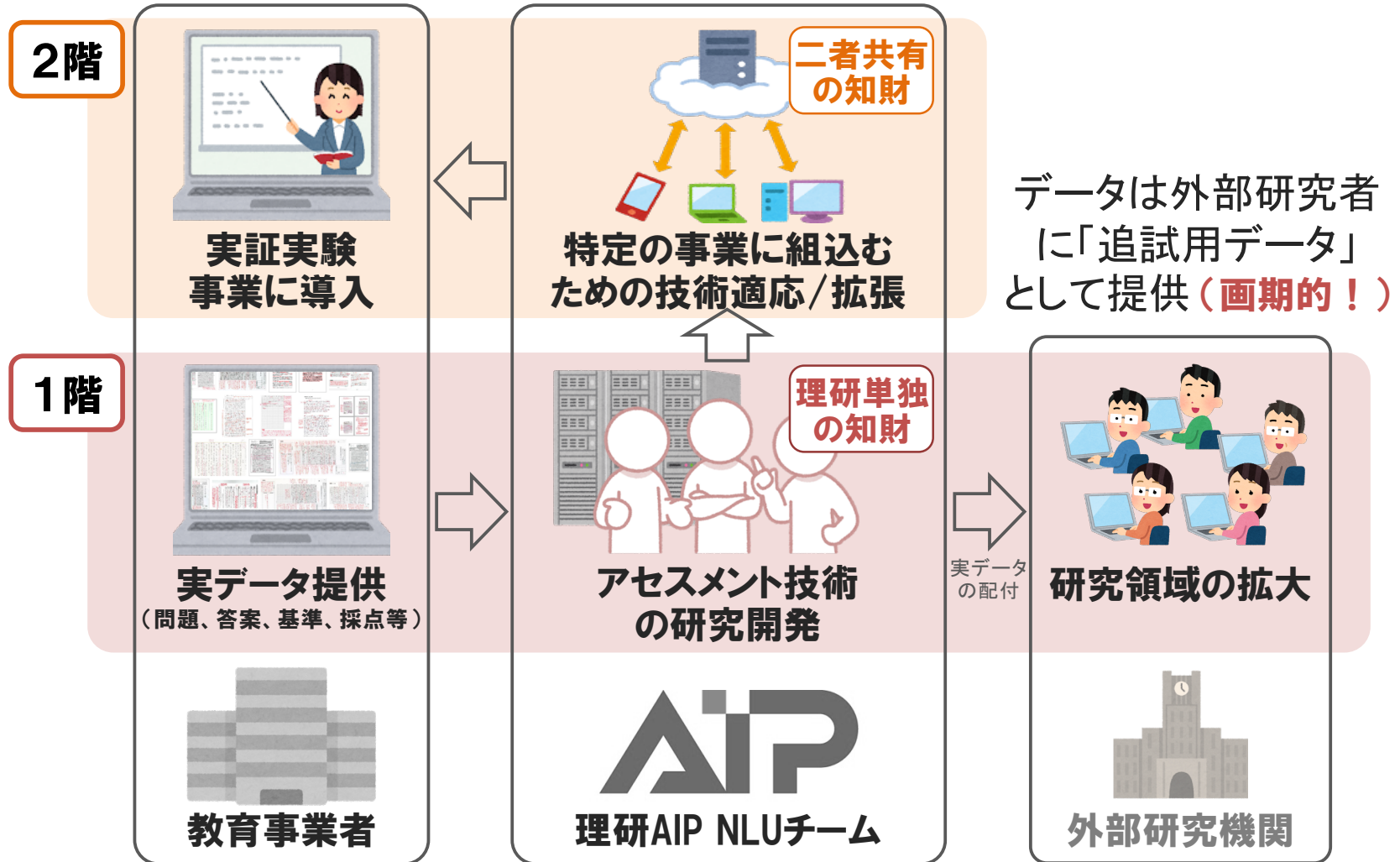


理研AIP NLUチーム

多様な教育事業者と2階建て連携



多様な教育事業者と2階建て連携





TOHOKU
UNIVERSITY

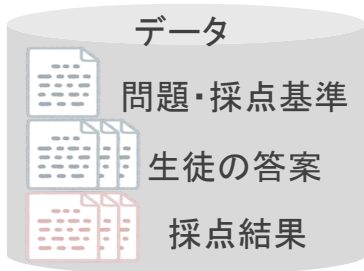
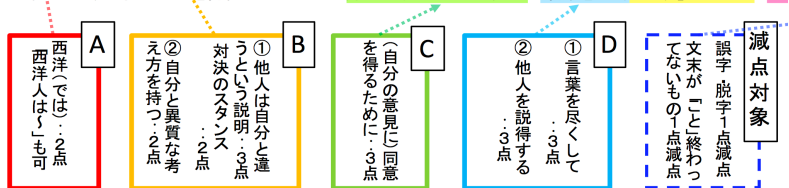


記述式答案の項目別自動採点

(Mizumoto+ BEA 2019; Funayama+ ACL-SRW 2020; Ou+ under review)

代々木ゼミナールデータ

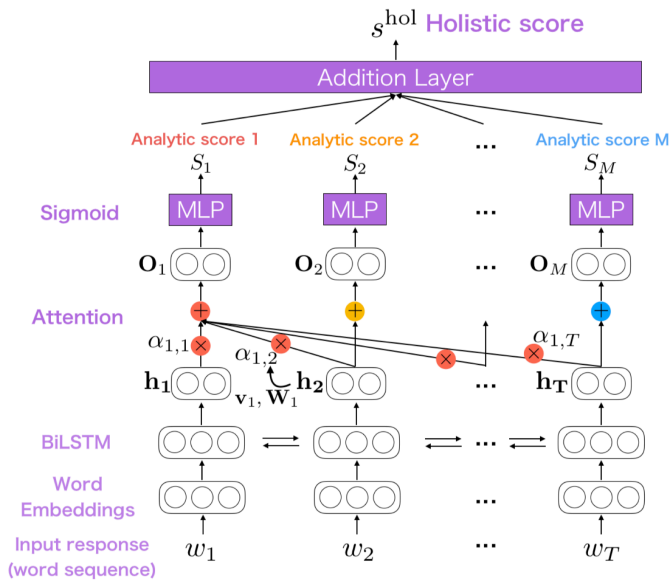
西洋文化の基底の「対決」は神対人間(宗教=契約)、人間対自然(科学=合理主義)、人間対人間(個人主義)とい
日本人のように余計なことを言わないのではなく、他人に分かってもらうために言葉を尽くす「対決」のスタンスが西



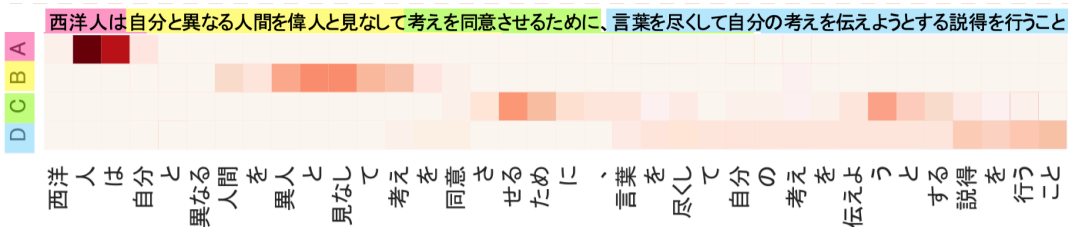
現代国語(高校)
記述24問×~2,000答案
(世界最大)

深層学習によるモデル化

200答案の訓練データで問題によっては人間と同等の精度



	Q1	Q2	Q3	Q4	Q5	Q6	Ave.
Analytic/Justification: 200 (QWK)							
SVR	.85	.87	.44	.77	.78	.71	.735
base	.92	.91	.57	.78	.83	.76	.794
+just.	.95	.92	.65	.80	.84	.78	.822
+hol.	.97	.94	.72	.82	.88	.83	.859
human	.96	.94	.76	.84	.82	.90	.873



理研記述問題採点データセット

2020年7月公開



大学共同利用機関法人 情報・システム研究機構
国立情報学研究所
National Institute of Informatics



情報学研究データリポジトリ

English

HOME データ一覧 研究成果一覧 ユーザフォーラム 組織 関連リンク 問い合わせ

> HOME > データ一覧 > 理研記述問題採点データセット

民間企業提供データ

- Yahoo!データセット
- 楽天データセット
- ニコニコデータセット
- リクルートデータセット
- クックパッドデータセット
- LIFULL HOME'Sデータセット
- 不満調査データセット
- Sansanデータセット
- インテージデータセット
- オリコンデータセット

理研記述問題採点データセット

国立情報学研究所が研究者提供データセットとして理化学研究所から受け入れたデータセットです。

2020/07/14 作成

データ概要

大学受験予備校や教育関連会社にて行われたテストの記述問題の解答に対し、理化学研究所において採点アノテーションを付与したデータセットです。今後も順次同様のデータの追加を予定しています。

1. 代々木ゼミナールデータ（2020年4月時点、9問各約2000件）

2014年～2015年に行われた高校生対象模擬試験の国語記述問題に採点アノテーションを付与したものです。全体点および採点項目ごとの部分点や、その根拠となる単語・フレーズがアノテーションされています。

データの詳細は[理化学研究所の「記述答案採点の自動化に関する研究」のページ](#)もご参照ください。

また本データを用いた研究成果として理化学研究所が発表した[国際会議ACL併催ワークショップBEA14の論文](#)もご参照ください。

さまざまな教育事業者との連携

理研AIPがハブとなってデータ・技術を集積・共有化



教育アセスメント×言語処理シンポジウム： 自動採点、英文添削、論述評価の可能性



主催：理研AIPセンター 自然言語理解チーム

共催：東北大学 データ駆動科学・AI教育研究センター、教育学研究科

2020年11月27日（金）

13:00 開会挨拶

「教育アセスメント×言語処理シンポジウム」の開催に当たって 乾健太郎（理研/東北大）

13:20 -15:10 技術紹介：教育アセスメント×言語処理の取り組み 司会：奥村学（東工大/理研）

1. 国語記述式答案自動採点の現状と実応用に向けて 舟山弘晃（東北大/理研）
2. 和文英訳問題の自動採点技術についての展望（仮） 松林優一郎（東北大/理研）
3. ライティング支援のための文法誤り訂正技術の現状と今後の展望 三田雅人（理研/東北大）
4. ライティング学習のための解説文生成 永田亮（甲南大/理研）
5. 論述の構造解析による反論スピーチへのフィードバック生成 内藤昭一（株式会社リコー/理研）

15:20-16:30 パネル討論：教育アセスメント×言語処理の可能性 司会：乾健太郎（理研/東北大）

1. 代ゼミ自動採点における今後の取り組み
代々木ゼミナール教育事業推進本部長 佐藤 雄太郎
2. 思考力の評価・育成と言語処理技術への期待
株式会社ベネッセコーポレーション 学校カンパニー GPS-Academic問題開発リーダー 伊藤 素江
3. 模範解答をなぞる英作文学習から卒業しよう！弱点を知り主体的なやり直しを促す支援ツールの開発
株式会社増進堂・受験研究社 NEXT LEARNING Labs主任研究員 岡田 健志
4. Classiにおける自然言語処理の取り組みと日本語の言語処理基礎研究に対する期待（仮）
Classi株式会社 データAI部長・データサイエンティスト 伊藤 徹郎
5. ディープラーニングで経営教育のディープラーニングを目指す
学校法人グロービス経営大学院教授兼GAI MERi所長 鈴木 健一
6. 全体討論・意見交換

我々がねらいたいもの

- 言語で説明できるタイプの問題領域をせめたい
 - DNN と記号推論の融合をどう設計するか
- マイルストーン(問題設計とデータ構築は大事)
 - 質問応答、ファクトチェック支援
 - NL Assessment 領域の開拓も魅力的