

aoba\_v2 bot:

# 多様な応答生成モジュールを統合した 雑談対話システム

---

長澤春希<sup>1</sup> 工藤慧音<sup>1</sup> 宮脇峻平<sup>1</sup> 有山知希<sup>1</sup>

成田風香<sup>1</sup> 岸波洋介<sup>1</sup> 佐藤志貴<sup>1</sup> 乾健太郎<sup>1,2</sup>

東北大学<sup>1</sup> 理化学研究所<sup>2</sup>

# 設計アイデア： 様々な応答生成機構を統合した雑談対話システム

評価指標

期待される応答

適するアーキテクチャ

自然性

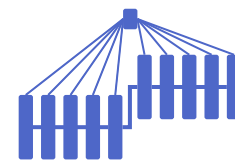
話題追従

話題提供

ユーザ発話・話題に即した自然な応答

話題に関する情報提供ができる応答

ニューラル  
応答生成機構



知識ベース  
応答生成機構



aoba\_v2 bot

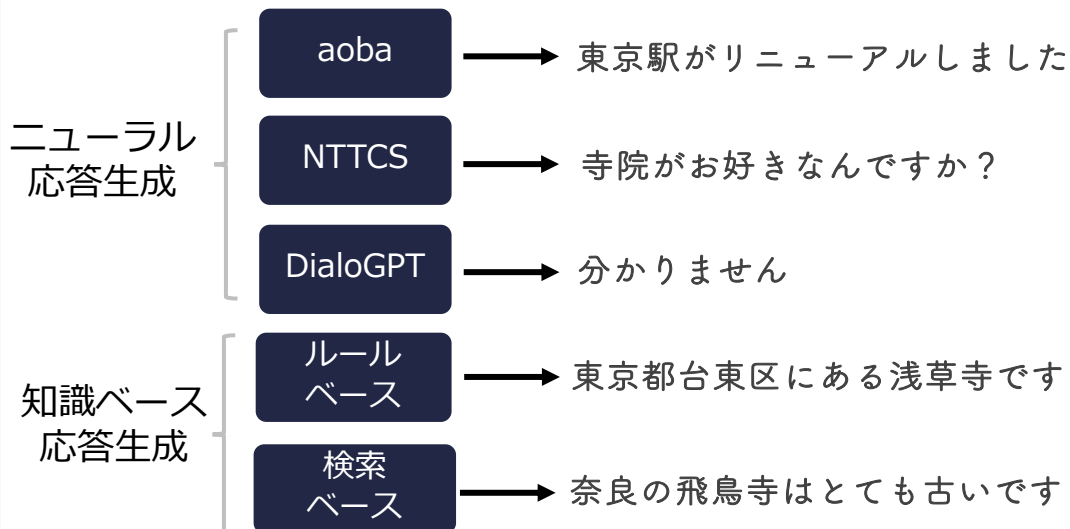
多様な応答生成機構で応答候補を生成し  
最適な応答を候補から選択

# TohokuNLP aoba\_v2の概要

Input

東京都内で最も古い寺院はどこ？

## 1. 応答生成機構



## 2. 応答リランキング機構

1. 東京都台東区にある浅草寺です
2. 東京駅がリニューアルしました
- ⋮

5観点で  
評価

- ・重複度
- ・継続度
- ・類似度
- ・情報量
- ・一貫性

Output

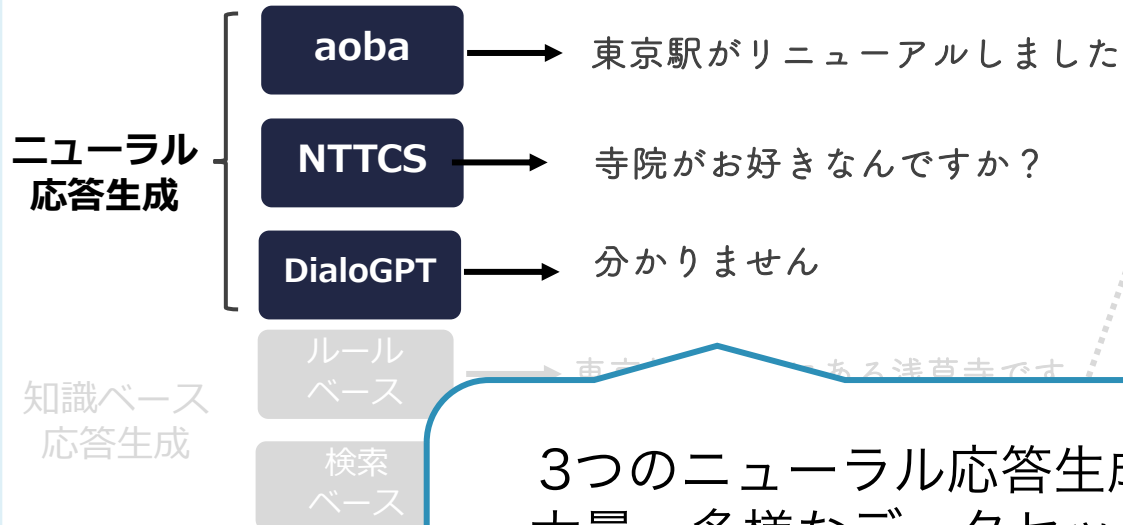
東京都台東区にある浅草寺です

# TohokuNLP aoba\_v2

## ニューラル応答生成機構

東京都内で最も古い寺院はどこ？

### 1. 応答生成機構



### 2. 応答リランキング機構

1. 東京都台東区にある浅草寺です
2. 東京駅がリニューアルしました

- ・重複度
- ・継続度

3つのニューラル応答生成機構を  
大量・多様なデータセットで訓練

↓  
「自然」な応答生成を目指す

# ニューラル応答生成機構

多様なテキストデータで学習し「自然」な応答を生成

3つのモデル

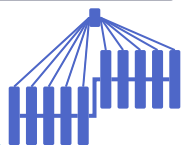
Pre-training

Further  
pre-training

Fine-tuning

**aoba**

[藤原+ '20]

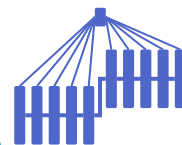


機構：  
Transformer

8.63億  
対話対

**NTTCS**

[Sugiyama+ '20]

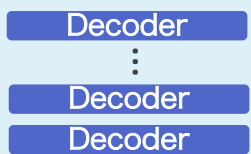


機構：  
Transformer

21億  
対話対  
公開モデル

**DialogPT**

[Zhang+ '20]



機構：  
GPT-2

1.47億  
対話対

12万  
対話対

20万  
対話対

# 学習データ：Pre-training

多様なテキストデータで学習し「自然」な応答を生成

3つのモデル

Pre-training

Further  
pre-training

Fine-tuning

aoba

[藤原+ '20]



機構：  
Transformer

8.63億  
対話対

大規模疑似対話データセット

Twitterリプライを対話とみなし、  
作成されたデータセット [藤原+'20]

NTTCS

[Sugiyama+ '20]



機構：  
Transformer

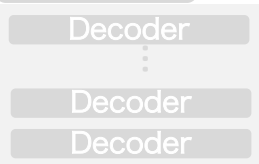
21億  
対話対  
公開モデル

対話対

対話対

DialoGPT

[Zhang+ '20]



機構：  
GPT-2

1.47億  
対話対

# 学習データ：Pre-training

多様なテキストデータで学習し「自然」な応答を生成

Further  
pre-training

Fine-tuning

## 翻訳データセット

機械翻訳モデル<sup>[kiyono + '20]</sup>を用い  
英語対話データを翻訳

DailyDialog

Blended Skill Talk

ConvAI2

Empathetic Dialogues

Wizard of Wikipedia

## 常識推論データセット

常識推論データセットから  
対話データを作成

京大常識推論データ

## 質問応答データセット

質問+答えから、疑似対話データを作成

JAQKET

解答可能性付き読解データ

12万  
対話対

20万  
対話対

# 学習データ：Pre-training

多様なテキストデータで学習し「自然」な応答を生成

3つのモデル

Pre-training

Further  
pre-training

Fine-tuning

aoba

[藤原+ '20]

機構：  
Transf

NTTCS

[Sugiyama+ '21]

機構：  
Transf

DialoGPT

[Zhang+ '21]

Decoder

Decoder

Decoder

機構：  
GPT-2

## 高品質データ

高品質なデータセットを [藤原+ '20; Sugiyama+ '21]  
組み合わせて作成

Twitter 疑似対話データ

対話破綻検出チャレンジデータ

JPersonaChat

JEmpathetic Dialogues

20万  
対話対

1.47億  
対話対



# TohokuNLP aoba\_v2

## 知識ベース応答生成機構

東京都内で最も古い寺院はどこ？

### 1. 応答生成機構

ニューラル  
応答生成

ao  
NT  
Dialog

知識ベースを用いることで  
「情報」を提供できる応答の生成を目指す

知識ベース  
応答生成

ルール  
ベース

→ 東京都台東区にある浅草寺です

検索  
ベース

→ 奈良の飛鳥寺はとても古いです

### 2. 応答リランキング機構

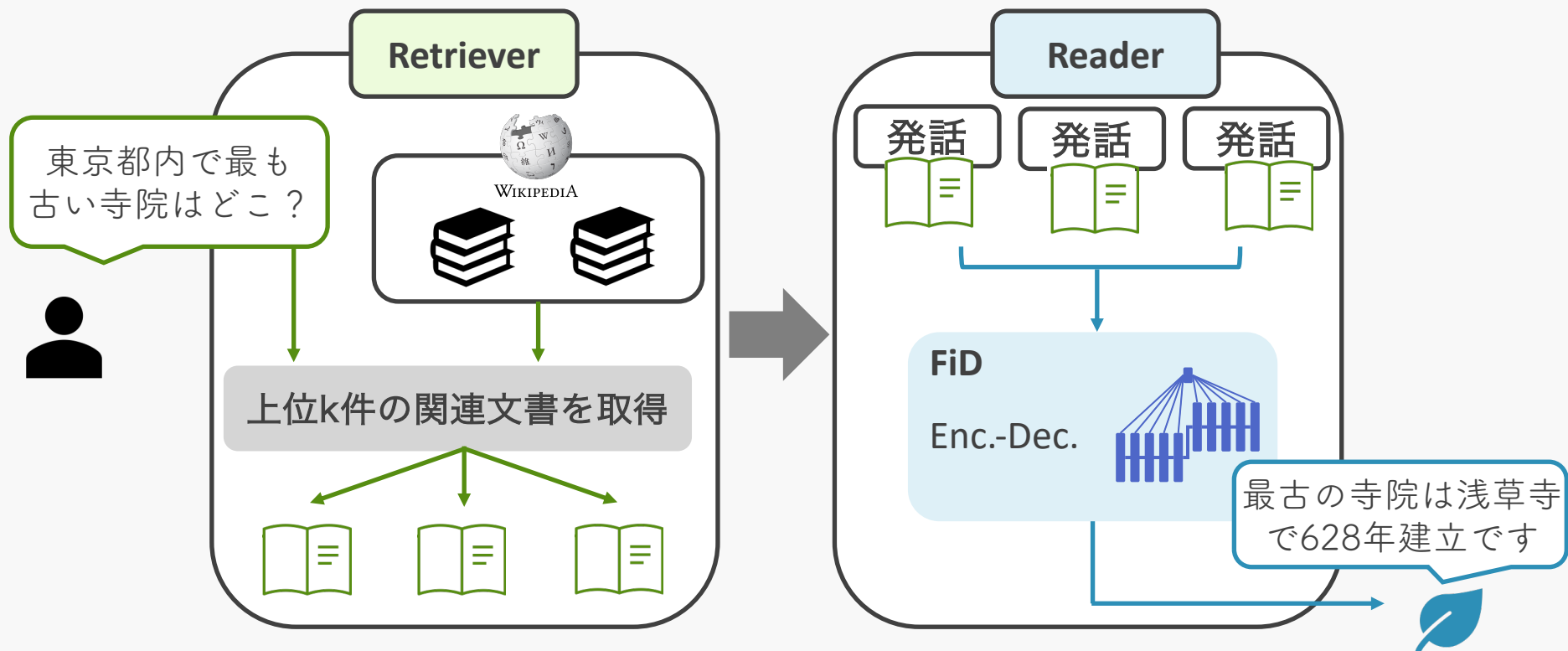
ある浅草寺です  
アルしました

- ・重複度
- ・継続度
- ・類似度
- ・情報量
- ・一貫性

東京都台東区にある浅草寺です

# 検索ベース応答生成モデル

- ① **Retriever** で、発話と関連する文書をWikipediaから検索
- ② **Reader** で、発話と①で検索された文書を使い、応答を生成



# TohokuNLP aoba\_v2

## 応答リランキング機構

東京都内で最も古い寺院はどこ？

### 1. 応答生成機構

aoba

東京駅がリニューアルしました

藤原らが提案した4つの手法

+

新たに提案する「一貫性」

↓

最適な応答を選択

### 2. 応答リランキング機構

1. 東京都台東区にある浅草寺です
2. 東京駅がリニューアルしました

⋮

5観点で  
評価

- ・重複度
- ・継続度
- ・類似度
- ・情報量
- ・一貫性

Output

東京都台東区にある浅草寺です

# 応答リランキング機構

生成された発話候補から**自然で適切な情報を提供**している応答を選択

東京のシンボリック的存在です。

昨年  
のモデル  
で使った  
指標

- **継続度**: 直前の発話に対し自然な応答か  
スカイツリーは大好きだよ < 東京のシンボリック的存在です。
- **重複度**: 同じ語句を繰り返していないか  
東京タワーは東京タワー < 東京のシンボリック的存在です。
- **類似度**: 直前の発話を復唱していないか  
東京タワーについてですね < 東京のシンボリック的存在です。
- **情報量**: 十分な情報量を含んでいるか  
それは分かりません。 < 東京のシンボリック的存在です。

追加した  
指標

- **一貫性**: 直前の自己発話やユーザ発話と矛盾しないか  
東京タワー嫌いなんですね < 東京のシンボリック的存在です。

↑ ↑ ↑ ↑ ↑  
**応答生成機構**

**応答候補**を生成

東京タワー好きなんだよね。  
東京タワーって知ってる？

# 新たに追加した指標：一貫性

生成された発話候補から**自然で適切な情報を提供している**応答を選択

東京のシンボルの存在です。

直前の自己発話やユーザ発話に  
**矛盾する**応答をBERTで検出

ex.) ユーザ発話：東京タワー好きなんだよね。

↑  
**矛盾検出**  
↓

応答候補：東京タワーが嫌いなんですね。

昨年のも  
で使った

在です。

在です。

在です。

それは分かりません。

のシンボルの存在です。

追加した  
指標

- **一貫性**: 直前の自己発話やユーザ発話と矛盾しないか

東京タワー嫌いなんですね < 東京のシンボルの存在です。

↑ ↑ ↑ ↑ ↑  
応答生成機構

応答候補を生成

東京タワー好きなんだよね。  
東京タワーって知ってる？

# システム対話例 (システム全体)

ユーザ



洋楽は好きですか?

洋楽はそんなに聞かないですね... 英語が難しいです。

英語は難しいですよね。でも、外国のアーティストの曲はかっこいいと思います。

そうですね。友人にも洋楽好きな人がいるのですが、彼女の英語力も高いので、英語をちゃんと勉強する必要がありそうです...

彼女は英語が得意なんですか?

そうですね。僕とは違って発音がきれいです。

aoba\_v2



# システム対話例 (モデル別)

マリオって知ってますか?



任天堂が発売したコンピュータ  
ゲームソフトのシリーズ、  
マリオシリーズに出てくる架空の  
キャラクターみたいですよ。



検索ベース  
応答生成モデル

最近、音楽を聴くのが  
趣味なんです!



今は、どんな音楽を  
聴いていますか?

どんな音楽が好きですか?  
好きなアーティストはいますか?

どんな音楽を聴きますか?  
僕は洋楽が好きですね。

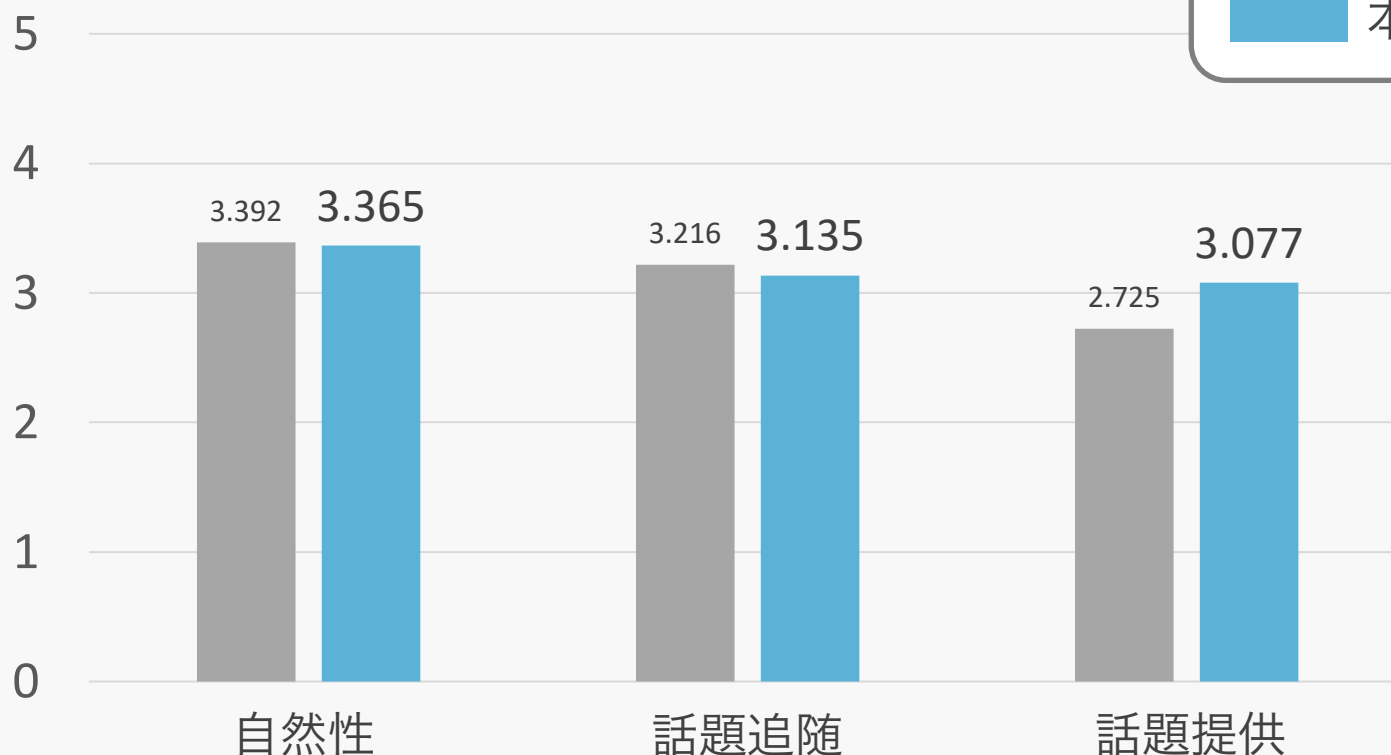
DialogPT

NTTCS

aoba

# 予選評価（スコア）

対話したワーカによる5段階評価の平均



- ✓概ね自然な応答生成が実現できた
- ✓話題提供に関し、大幅なスコアの改善が見られた



# 予選評価（ワーカの声）

## 良かった点

- 『自然な感じで楽しかったです。』
- 『とても知的な方だと感じました。』

自然で情報量の多い応答を生成できた

## 悪かった点

- 『会話の切り替えがいささかぎごちなかった』
- 『唐突に話題が切り替る傾向があると感じました。』

話題転換時の応答が不自然  
話題転換を検出する仕組みが必要

# まとめ

- 異なる仕組みによって応答を生成する複数のモジュールを統合した対話システムを構築
- 応答候補を5つの指標でリランキング
- 評価結果から、構築した対話システムが**自然で情報量の多い応答**を実現できたことを確認

## 課題・今後の展望

- 話題転換時の応答が不自然  
→話題転換を検出する仕組みの追加

# 参考文献

- 藤原吏生, 岸波洋介, 今野颯人, 佐藤志貴, 佐藤汰亮, 宮脇峻平, 加藤拓真, 鈴木潤, 乾健太郎. ILYS aoba bot: 大規模ニューラル応答生成モデルとルールベースを統合した雑談対話システム. 第90回人工知能学会言語・音声理解と対話処理研究会(第11回対話システムシンポジウム), pp. 110–115, 2020.
- Hiroaki Sugiyama, Masahiro Mizukami, Tsunehiro Arimoto, Hiromi Narimatsu, Yuya Chiba, Hideharu Nakajima, and Toyomi Meguro. Empirical analysis of training strategies of transformer-based japanese chat systems. In arXiv:2109.05217, 2021.
- Yizhe Zhang, Siqi Sun, Michel Galley, Yen-Chun Chen, Chris Brockett, Xiang Gao, Jianfeng Gao, Jingjing Liu, and Bill Dolan. DialoGPT: Large-scale generative pretraining for conversational response generation. In ACL, pp. 270–278, 2020.
- Shun Kiyono, Takumi Ito, Ryuto Konno, Makoto Morishita, and Jun Suzuki. Tohoku-AIP-NTT at WMT 2020 news translation task. In WMT, pp. 145–155, 2020.
- 東中竜一郎, 船越孝太郎, 小林優佳, 稲葉通将.  
対話破綻検出チャレンジ.  
第75回言語・音声理解と対話処理研究会(第6回対話システムシンポジウム)
- (松田耕史, 鈴木正敏, 乾健太郎).  
Wikidataからの遠距離教師あり学習に基づく大規模関係知識獲得.  
言語処理学会第25回年次大会(NLP2019))

# Appendix

---

# ルールベース応答生成モデル

知識を問うような発話に対し

知識ベースと応答テンプレートを使用して発話を生成

**OUTPUT** 東京タワーは東京都にあります。東京都港区にある電波塔です。

知識ベース

主語：東京タワー

目的語：東京都

述語：location

テキスト：東京タワーは(東京都港区)  
にある電波塔である

応答テンプレート

質問フレーズ：どこ

述語：location

テンプレート：(主語)は  
(目的語)にあります

名詞句抽出

質問フレーズ抽出



東京タワーってどこにあるの？

# 各種モデルのパラメータ概数

- aoba: 約4.8億
- NTTCS: 約16億
- DialoGPT: 約3.5億
- FiD(検索ベース応答生成): 約2億