# What Did We Learn from the Dialogue System Live Competition?

特　集「AI コンペティション」

# 対話システムライブコンペティションから何が得られたか

## What Did We Learn from the Dialogue System Live Competition?

東中　竜一郎[†1]　日本電信電話株式会社 NTT メディアインテリジェンス研究所
Ryuichiro Higashinaka　NTT Media Intelligence Laboratories, NTT Corporation
ryuichiro.higashinaka.tp@hco.ntt.co.jp

船越　孝太郎[†2]　京都大学
Kotaro Funakoshi　Kyoto University
funakoshi.k@i.kyoto-u.ac.jp

稲葉　通将　電気通信大学
Michimasa Inaba　The University of Electro-Communications
m-inaba@uec.ac.jp

角森　唯子　（株）NTT ドコモ
Yuiko Tsunomori　NTT DOCOMO INC.
yuiko.tsunomori.fc@nttdocomo.com

高橋　哲朗　（株）富士通研究所
Tetsuro Takahashi　Fujitsu Laboratories, LTD.
takahashi.tet@jp.fujitsu.com

赤間　怜奈　東北大学，理化学研究所
Reina Akama　Tohoku University / RIKEN AIP
reina.a@ecei.tohoku.ac.jp

宇佐美　まゆみ　国立国語研究所
Mayumi Usami　National Institute for Japanese Language and Linguistics
usamima@ninjal.ac.jp

川端　良子　（同　　上）
Yoshiko Kawabata
kawabata@ninjal.ac.jp

水上　雅博　日本電信電話株式会社 NTT コミュニケーション科学基礎研究所
Masahiro Mizukami　NTT Communication Science Laboratories
masahiro.mizukami.df@hco.ntt.co.jp

**Keywords:** dialogue system, chat-oriented dialogue system, evaluation, competition.

## 1. は　じ　め　に

工学の分野では多くのコンペティションが開催され，それによって進展がもたらされてきた．そして，それは対話システムにとっても同じである．特に，近年は対話システムに関わるコンペティションが多く開催されている [東中 19a].

表 1 は対話システムに関する主なコンペティションをまとめたものである．軸は二つある．一つは対象とする対話システムの種類である．対話システムは，所定のタスクを遂行することを目的とするタスク指向型対話システムと，所定のタスク遂行が主な目的ではない非タスク指向型対話システム（雑談対話システム）に大別されるが [中野 15], 表ではこの分類を用いている．もう一つは，評価形態である．対話システムの評価の仕方には，オフライン評価とオンライン評価がある．前者は，固定的なデータセットを対象に，対話システムの特定のモジュー

†1 2020 年 4 月より，名古屋大学大学院情報学研究科.
†2 2020 年 4 月より，東京工業大学科学技術創成研究院未来産業技術研究所.

表1 対話システムに関する主なコンペティション

| | オフライン評価 | オンライン評価 |
|---|---|---|
| タスク指向型対話システム | ・Dialogue state tracking challenge（DSTC） | ・DiaLeague<br>・Spoken Dialogue Challenge |
| 非タスク指向型対話システム（雑談対話システム） | ・対話破綻検出チャレンジ（DBDC）<br>・NTCIR Short Text Conversation（STC） | ・ローブナー賞<br>・Amazon Alexa Prize<br>・Conversational Intelligence Challenge<br>・対話システムライブコンペティション |

ル（例えば，発話理解や発話選択）の性能を測る．例えば，対話文脈とその文脈で得られるべきユーザ意図が対になったデータセットが与えられ，対話文脈からユーザ意図を推定する精度で競ったり [Henderson 14]，対話文脈と複数のシステム発話候補が対になったデータセットが与えられ，所定の文脈で複数の候補から妥当なシステム発話を選択する精度を競ったりする [Hori 19]．後者のオンライン評価は，対話システムの End-to-End の性能を測るもので，ユーザとシステムが実際に対話を行い，主に主観により対話の質を評価する．

オフライン評価は，パラメータをいくつも振って最適化を図ることができるなど，多くのアルゴリズムを効率的に試すことができるというメリットがある．しかし，切り取られた一部の対話文脈しか扱うことができないというデメリットがある．オンライン評価はその裏返しで，ユーザとシステムの発話によってつくられていくダイナミックな文脈を扱うことができるというメリットがある．しかし，人手による評価が必要となり，コストがかかるというデメリットがある．

著者らが提案・実施してきた，対話システムライブコンペティション（以降，ライブコンペ）は，非タスク指向型対話システムを対象としたオンライン評価を行うコンペティションである [東中 18, 東中 19b, Higashinaka 19c]．これまでにも，表1で同じ象限に含まれるコンペティションはなかったわけではない．しかし，ライブコンペにはこれまでのものとは異なるポイントがある．それは，対話のダイナミックさにフォーカスしていること，そして，対話システム研究のコミュニティの問題意識に働きかけるということである．

先ほど述べたとおり，対話とは発話ごとに状況が刻々と変化していくダイナミックなプロセスである．そのダイナミックさを対話システムの研究者・開発者が体感できるコンペティションをつくろうと考えた．具体的には，対話システムと人間の話者の対話を，それがあたかもライブコンサートであるかのように，研究者・開発者全員でライブで鑑賞し，その良し悪しを評価するということを考えた．それがライブコンペという企画である．

全員で同じ対話を鑑賞するということは，対話システム研究にとって大きな意味がある．それは，現状の問題点をコミュニティ全員で共有できるということだ．これにより，対話システム研究が抱える多くの課題の中で本当に着手すべき課題にコミュニティが一丸となって取り組める可能性が高まる．

本稿では，これまでに二度実施してきたライブコンペの仕様，結果や得られた問題意識について述べる．また，ライブコンペに関連するイベントとして開催した，対話システムライブコンペ講習会と日本語教育学会におけるパネルセッション [宇佐美 19] についても触れる．

## 2. 対話システムライブコンペティション1

第1回対話システムライブコンペティション（ライブコンペ1）は第9回対話システムシンポジウム（2018年11月20～21日）の1セッションとして開催された．対話システムシンポジウムとは，本学会の言語・音声理解と対話処理研究会（SIG-SLUD）が2010年から年に一度開催している，国内の対話システム関係者が一堂に会するイベントであり，ライブコンペの趣旨に鑑みて最適な場である．以下，ライブコンペ1について紹介する．詳細はオーバビュー原稿 [東中 18, Higashinaka 19c] を参考にされたい．

### 2·1 仕　　様

**§1 対象とする対話システム**

対象とする対話システムは非タスク指向型対話システムとした．タスク指向型対話システムはバックエンドの知識源（レストランのデータベースなど）が必要となり準備コストが高いこと，また，非タスク指向型対話システムは基礎的研究の段階であり，企業・大学問わず参加しやすいと考えられることから，このように設定した．

**§2 評価尺度**

「どのくらいまた話したいと思うか」という一つの軸で評価することにした．非タスク指向型対話システムに望まれる要素として，少なくとも対話を継続することが重要だと考えたためである．5段階のリッカート尺度で評価する．これは，Amazon Alexa Prize[*1] と同様である．

**§3 対話システムの要件**

Telegram[*2] と呼ばれるメッセージングプラットフォーム上でボットとして動作することを要件とした．これは，Conversational Intelligence Challenge [Dinan 19] と同様である．Telegram は API が充実しており，ボットの構築が容易である [東中 20]．

システムはテキストのみで対話することとし，絵文字，顔文字，スタンプの使用は不可とした．また，システムは1ターンで1発話のみを行うこととした．対話はシステム発話から始まり，16ターン目でシステムから対話

*1 https://developer.amazon.com/alexaprize
*2 https://telegram.org/

表 2　ライブコンペ 1 における各システムの手法.
方式はルールベース，抽出ベース，生成ベースの中から，知識源は，大規模テキストデータの利用，知識ベースの利用，対話データの利用の中から該当するものを参加チームが選択した．学習手法については，機械学習を用いていない場合は N/A としている

| 順位 | チーム名 | 方　式 | | | 知識源 | | | 学習手法 |
|---|---|---|---|---|---|---|---|---|
| | | ルール | 抽　出 | 生　成 | テキスト | KB | 対　話 | |
| 1 | NTTdocomo | ✓ | | ✓ | ✓ | | | Other |
| 2 | NTTCS | ✓ | | | ✓ | ✓ | ✓ | CRF，SVM |
| 3 | teamzunko | ✓ | | ✓ | ✓ | ✓ | ✓ | RNN |
| 4 | IRS | | ✓ | | | | ✓ | N/A |
| 5 | TEAM1 | ✓ | ✓ | | | | ✓ | N/A |
| 6 | TEAM2 | ✓ | | ✓ | ✓ | | | CNN，RNN |
| 7 | TEAM3 | ✓ | ✓ | | ✓ | | ✓ | CNN |
| 8 | TEAM4 | ✓ | ✓ | | | | ✓ | RNN |
| 9 | RSL | ✓ | | | | | | N/A |
| 10 | TEAM5 | | | ✓ | | | ✓ | RNN |
| 11 | TEAM6 | | | ✓ | | | ✓ | Transformer |

を終えることとした．16 ターンという対話の長さについては，現状の対話システムの性能に鑑み設定した．

§ 4　実施の形態

現状の問題点をコミュニティで共有するためには，現時点でなるべく高性能な対話システムを選定する必要がある．そこで，予選を実施することにした．予選はクラウドソーシング*3 を用いて行い，勝ち抜いた上位のシステムが対話システムシンポジウムにおけるライブイベントで人間の話者との対話を披露できることにした．

§ 5　情報公開のポリシー

企業からの参加をしやすくするために，予選を通過したチームのみ，組織名などの情報を公開することにした．その他のチームについては希望者のみ公開することにした．

## 2·2　予　　選

2018 年 9 月 30 日のエントリ締切までに，12 チーム（オーガナイザの 1 チームを含む）がエントリを行った．その後，疎通確認ができた 11 チームについてクラウドソーシングによる評価を実施した．クラウドワーカ（最大 30 名*4）が各対話システムと対話し，「どのくらいまた話したいと思うか」について評価した．平均スコアの上位 3 チームがライブイベントに進出した．予選のスコアは図 1 のとおりである．ここで，スコアは 1 が最も良く，5 が最も悪いことに注意されたい．TEAM 1-6 は組織名非公開のチームを示す．

予選に勝ち残ったチームは NTTdocomo，NTTCS，および，teamzunko であり，それぞれ状態遷移ベースで旅行についての雑談が可能なシステム，大規模テキストから得られた知識を用いオープンドメインの雑談を行うシステム，複数のデータベースおよび深層学習に基づく発話生成を組み合わせたシステムであった．予選を通過したチームはいずれも，ルール，生成，知識源などの組合せを行っており，対話データに基づく生成モデルのみを利用したチームは順位が低かった．参加チームの手法については表 2 を参照されたい．

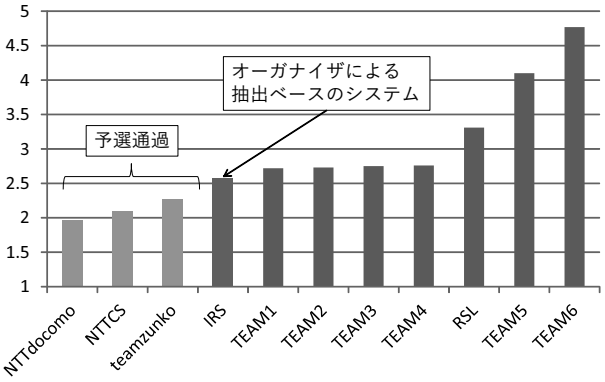



図 1　ライブコンペ 1 の予選の結果．スコアは低いほうが良い

## 2·3　ライブイベント

対話システムシンポジウムでは，予選上位 3 チームが人間の対話者と二度ずつ対話した．対話者は本学会言語・音声理解と対話処理（SLUD）研究会主査およ

*3 https://crowdworks.jp/

*4　30 人のクラウドワーカに評価を割り振ったが，クラウドソーシングの特性上，評価に問題があるケースがあり，それらを排除した結果 30 人未満の評価になった場合があったため，最大 30 名と記載している．

図 2　システムと対話者が対話する様子

図 3　参加者が対話を鑑賞・評価している様子

びオーガナイザが選定した COI の認められない対話研究者である．対話順はくじ引きで決定した．図 2 はシステムと対話者が対話している様子である．図 3 は参加者が対話を鑑賞・評価している様子である．これらの写真からもわかるとおり，大勢（100 人超）の参加者が同じ対話をライブで鑑賞・評価した．評価軸は，予選と同様，「どのくらいまた話したいと思うか」である．参加者は，16 ターンからなる一連の対話の後すぐに Google Form を用いて評価値を投入した．この際，発話ごとに Good，Bad，どちらでもないの 3 段階の評価も投入した．

ライブイベントにおける評価の結果は，NTTCS が 2.81（1 回 目：2.69，2 回 目：2.93） で 1 位，NTTdocomo が 3.06 点（1 回目：3.18，2 回目：2.95）で 2 位，teamzunko が 3.48 点（1 回目：3.58，2 回目：3.40）で 3 位であった．いずれのスコアも約 100 名の評価値の平均である．NTTCS のシステム [杉山 18] は，旅行というドメインにユーザを引き込み，細かく定義された状態遷移に従って，かみ合った雑談を実現していた点が評価された．その他のシステムは，オープンドメインを志向していたこともあり，ユーザの自由発話に対応できない局面も見られ，対話破綻 [Higashinaka 16] の状況に陥ることがままあった．図 4 は NTTCS のシステムの対話（1 回目）である．少し長いがどのような対話が行われたかが具体的に理解しやすいと思われるので全文を掲載する．

| 話者 | 発話 | ○ | △ | × |
|---|---|---|---|---|
| U | /start | | | |
| S | こんにちは！お名前を伺ってもいいですか？ | 93.8% | 5.2% | 1.0% |
| U | 伝といいます | | | |
| S | 伝さんとおっしゃるんですね。私はリコっていいます．よろしくお願いします．伝さんは何県にお住まいなんですか？ | 95.8% | 3.2% | 1.0% |
| U | いまは千葉県です | | | |
| S | ふむふむ。千葉県ですか。千葉いいですね。行ってみたいです。東京ディズニーランドとか有名ですよね？ | 93.8% | 4.1% | 2.1% |
| U | あんまり千葉県と認識されてないですけどね | | | |
| S | あ、そうなんですか。前に行った人がおすすめって言ってたので、実際どうなのかなあと。千葉県で伝さんおすすめの観光地ってどこかあります？ | 75.0% | 16.7% | 8.3% |
| U | 千葉県、うーん、ないですね | | | |
| S | まあそんなもんですよね。千葉県だと、清水公園とかも有名ですよね？ | 76.8% | 13.7% | 9.5% |
| U | 清水公園知らないですね。どこですか | | | |
| S | ふむふむ。清水公園って行ったことあります？ | 14.0% | 10.7% | 75.3% |
| U | ないです | | | |
| S | あ、そうなんですか。前に行った人がおすすめって言ってたので、実際どうなのかなあと。ところで、あ、こないだ三重県の番組見たんですけど、伊勢神宮とかおすすめですよ。行かれたことあります？ | 34.7% | 42.1% | 23.2% |
| U | あります | | | |
| S | 私は神社とか好きなんですけど、伊勢神宮の神社もいいですよ。伝さんはどう思われます？ | 44.2% | 41.1% | 14.7% |
| U | すごく大きかったですね。いついったんですか？ | | | |
| S | なるほど。。今度行った時の参考にさせてもらいます。そういえば、伝さんは最近旅行で何県に行かれました？ | 31.9% | 23.4% | 44.7% |
| U | 最近は登山で、よく長野県とか行きますね | | | |
| S | なるほど。長野県ですか。長野いいですね。行ってみたいです。神社といえば戸隠神社とかもいいらしいですね。伝さんは行かれたことはあります？ | 53.7% | 31.6% | 14.7% |
| U | この前の８月に行きました | | | |
| S | 長野県に８月ってところがあるんですね。いいですね。８月の神社ってどうですか？ | 6.2% | 8.4% | 85.4% |
| U | 涼しくっていいですよ | | | |
| S | おー、それはぜひ行ってみたいです。そういえば、私奈良県出身なんですけど、法隆寺とかおすすめですよ。行かれたことあります？ | 51.0% | 31.3% | 17.7% |
| U | 子供のころ行きましたね | | | |
| S | ふむふむ。私いろんな建物を見るのも好きなんですけど、法隆寺もなかなかいいですよ。伝さんはどう思われます？ | 37.4% | 38.4% | 24.2% |
| U | 昔のことなので、忘れちゃいました | | | |
| S | ふむふむ。勉強になります。そういえば、伝さんはご出身はどちらなんですか？ | 31.1% | 30.0% | 38.9% |
| U | 大阪です | | | |
| S | あ、大阪ですか。大阪いいですね。行ってみたいです。あ、そろそろ終わりですね。お話ししていただいてありがとうございました！ | 62.6% | 27.5% | 9.9% |
| U | じゃ、さよなら | | | |
| S | ではでは。またお会いできるのを楽しみにしてます。 | 73.3% | 20.0% | 6.7% |

図 4　SLUD 研究会主査（当時）の伝 康晴氏と NTTCS のシステム（tripfreak）との対話．
U はユーザ，S はシステムを表す．○△×のカラムは，評価者が付与した発話ごとの評価値の割合（それぞれ Good，どちらでもない，Bad に対応）を示す．「/start」は対話の開始を表すコマンドである

### 2·4　課　　題

ライブで対話を鑑賞・評価し，対話システム研究者は何を課題と感じ取ったのだろうか．この調査のため，ライブイベントの事前と事後に，現状の対話システムの課題は何だと思うかについてのアンケートを取っている．このアンケートでは，対話システムの課題 30 項目（オーガナイザが対話研究者にヒアリングを行い列挙したもの．表 3 参照）のうち，特に重要だと思う 5 項目を選択してもらった．ライブイベントの前後で同じアンケートを行うことによって，ライブイベントが対話研究者にどのような問題意識をもたせたのかを見て取ることができる．

ライブイベント後のアンケートで最も重要だと思われた課題の上位 10 件は表 4 のとおりであった．各項目についてライブイベント前の順位も掲載している．ライブイベント前は「発話内容が文脈に沿っていない」が最も

表 3　対話システムの課題 30 項目.
課題は八つの観点から整理されている

| 観　点 | 課　題 | | | | |
|---|---|---|---|---|---|
| 1. 対話の意義 | 対話の目的が不明確 | システムの意図が不明確 | 対話から有益な情報が得られない | 対話に面白さがない | |
| 2. 話題 | 広い話題に対応できない | 適切な話題遷移ができない | | | |
| 3. 知識 | 常識が欠如している | 社会性が欠如している | 環境（対話している季節や時間など）の理解が欠如している | 事実と異なる情報を伝える | 発話に含まれる情報が古い |
| 4. ユーザ適応・個性 | ユーザに応じた応答ができていない | 会話を進めてもシステムとの関係性が変わらない | 過去のやり取りを活用できていない | キャラクタ性がない／一貫性に欠ける | |
| 5. 発話理解 | 発話内容の理解能力が低い／浅い | 感情の理解能力が低い | 質問に答える能力が低い | | |
| 6. 発話生成 | 日本語として不自然な文を生成する | 発話内容が不明瞭／不明確 | 発話内容が文脈に沿っていない | 同じ内容を繰り返し生成する | 自身の発話と矛盾した発話を生成する |
| 7. その他の機能 | 適切な主導権の交代ができない | 対話破綻から適切にリカバリできない | どのような会話をシステムができるのかが不明確 | 言語以外の入出力に対応できない | |
| 8. ビジネス・開発面 | 構築コストが高い | ビジネスモデルが不明確 | 品質評価が困難 | | |

表 4　ライブコンペ 1 後に参加者が重要と考えた対話システムの課題上位 10 件.
ランキングの変化とは，ライブイベントの前後で課題の順位がどのように変動したかを示している

| 順位 | 課題 | ランキングの変化 | |
|---|---|---|---|
| 1 | 【発話理解】発話内容の理解能力が低い／浅い | 2 => 1 | ↗ |
| 2 | 【話題】適切な話題遷移ができない | 8 => 2 | ↗ |
| 3 | 【発話理解】質問に答える能力が低い | 13 => 3 | ↗ |
| 4 | 【発話生成】発話内容が文脈に沿っていない | 1 => 4 | ↘ |
| 5 | 【対話の意義】システムの意図が不明確 | 18 => 5 | ↗ |
| 6 | 【ユーザ適応・個性】過去のやり取りを活用できていない | 11 => 6 | ↗ |
| 7 | 【対話の意義】対話の目的が不明確 | 7 => 7 | → |
| 8 | 【発話生成】同じ内容を繰り返し生成する | 19 => 8 | ↗ |
| 9 | 【ユーザ適応・個性】ユーザに応じた応答ができていない | 6 => 9 | ↘ |
| 10 | 【話題】広い話題に対応できない | 5 => 10 | ↘ |

重要な課題と認識されていたが, ライブイベント後は「発話内容の理解能力が低い／浅い」が 1 位となっている. その他,「適切な話題遷移ができない」や「質問に答える能力が低い」が上位に見られ，対話における基本的な理解・応答性能についてまだまだであるとの認識が共有されたと考えられる.

## 3. 対話システムライブコンペティション 2

ライブコンペ 1 では対話システムにおけるさまざまな課題が明らかになったこともあり，第 2 回（ライブコンペ 2）を企画・実施した．ライブコンペ 2 は第 10 回対話システムシンポジウム（2019 年 12 月 2 ～ 3 日開催）の 1 セッション（初日の午後）として開催された.

ライブコンペ 1 で志向していた雑談対話システムはとにかく話し続けたくなるというものであったが，対話における人間らしさ（特に，人間はシチュエーションによって言語運用が大きく異なる点に対する考慮）を置き忘れているのでは，という指摘が人文系の研究者からあり，特定のシチュエーションにおいて人間らしい対話ができたかどうかで評価する「シチュエーショントラック」を新設した．これに伴い，従来のライブコンペ 1 の内容は「オープントラック」という名前とした．以下，ライブコンペ 2 について述べる．詳細はオーバビュー [東中 19b] を参照されたい.

### 3·1　仕　　様

オープントラックの仕様は前回と同じとした．新設したシチュエーショントラックの仕様は基本的にオープントラックのもの（ライブコンペ 1 の仕様）と同じであるが，以下の点において違いがある.

- 評価尺度として「どれくらい（シチュエーションに適した）人らしい会話か」を用いる．評価は 5 段階のリッカート尺度で行う．シチュエーションとしては，図 5 に示すものを用いる.
- テキスト以外に，絵文字・顔文字も使用可能とする．絵文字・顔文字は人間どうしのテキストチャットではよく用いられており，使用不可とすることは人間らしい対話を志向するうえで自然ではないと考えたからである.

シチュエーショントラックでは，人間どうしの関係性，対話が発生する場所・時間，話題などが設定されており，その中で人間どうしが行うような対話を実現できるかを競う.

| | |
|---|---|
| システム | 名前:田中アイ (女)/アキラ (男), 年齢:20〜30 代, 職業:会社員 |
| ユーザ | 名前:鈴木ユウコ (女)/ユウキ (男), 年齢:20〜30 代, 職業:会社員 |
| 話者の関係 | 同性同士, 学生時代の友人関係 |
| 場所・時間 | 自宅, 暇な時間 |
| 話題 | 一番印象に残った旅行・場所 |

田中と鈴木は，学生時代，仲の良い友人同士であった．2人とも大学を卒業して会社員になってからはときどき食事に行ったりしていたが，ここ2，3年は会う機会も連絡をとることもなくなっていた．ある日，田中が自宅でのんびり過ごしていると，鈴木からテキストメッセージが送られて来た．鈴木も家で暇にしていたらしく，ふと気になって連絡をくれたらしい．久しぶりにお互いの近況報告をする中で，最近出かけた場所などが話題になった．田中「ところで，これまで行ったところで一番印象に残った場所ってどこ？」

図5 シチュエーショントラックにおいて設定したシチュエーション

### 3·2 予　　選

エントリの締切は2019年10月1日であった．この期日までに，オープントラックには9チームが，シチュエーショントラックには7チームがエントリした（それぞれオーガナイザのチームを含む）．オープントラックは9チームすべてが疎通確認でき，クラウドソーシングによる予選に進んだ．シチュエーショントラックでは，1チームの疎通確認ができなかったため，6チームがクラウドソーシングによる予選に進んだ．

ライブコンペ1では，最大30名のクラウドワーカにより評価を行ったが，対話システムの評価は主観要素が強く，より多くの評価者で評価することが望ましいと考えたことから最大50名による評価を行った．

予選の結果は図6および図7のとおりであった．なお，ライブコンペ1では，スコア1を「最も良い」としていたが，わかりにくいとの声もあり，ライブコンペ2では，5を「最も良い」，1を「最も悪い」としている点に注意されたい．

オープントラックでは，昨年優勝したNTTCSのシステムが3位となった．その上にTokoChanTeamとUEC（Tripia）が同スコアで入った．これらの3チームは他のチームと点差が見られることから予選通過とした．

シチュエーショントラックでは，OUHRIが1位となり，次いでオーガナイザによるルールベースのベースライン，そして，NTTCS[*5]という順位となった．これらの3チームを予選通過とした．

各システムの特徴を分析したところ，知識源に多少の違いはあれど，すべてのチームがルールベースのシステムであった．これは，ライブコンペ1でルールベースのNTTCSが1位となったこと，また，対話をなるべくかみ合うものにするために，手作業によるチューニングが必要と考えたためだと考えられる．なお，オープントラックにおいては抽出ベースや生成ベースのシステムのエントリも見られたが，いずれも予選を通過できなかった．現状で，対話として成立するためには，ルールベースのシステムが中心といえる．参加チームの手法については表5を参照されたい．

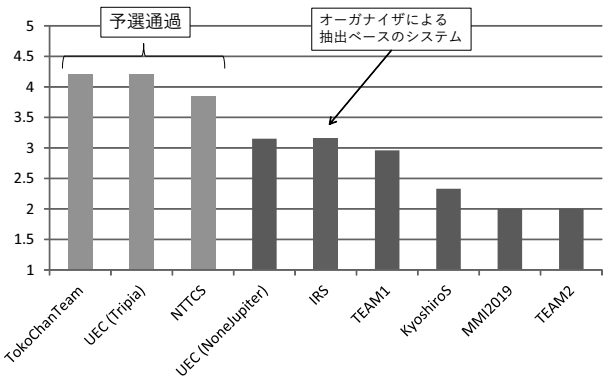



図6 ライブコンペ2（オープントラック）の予選の結果．スコアが大きいほうが良い

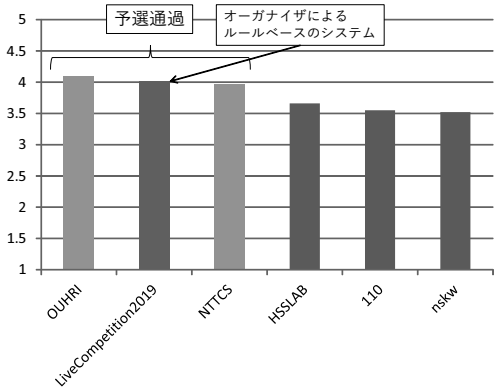



図7 ライブコンペ2（シチュエーショントラック）の予選の結果．スコアが大きいほうが良い

### 3·3 ライブイベント

ライブイベントでは予選を通過したシステムが人間の話者と対話し，会場の参加者が評価を行った．シチュエーショントラックが新設された影響により，各システムと人間の話者の対話は一度ずつとなった．

まず，オープントラックの対話がライブで披露され，その後，シチュエーショントラックの対話が披露された．参加者は，対話を鑑賞し，対話システムの各発話および対話全体についてGoogle Formを用いて評価した．評価軸は，オープントラックでは「どのくらいまた話したいと思うか」であり，シチュエーショントラックでは「どれくらい（シチュエーションに適した）人らしい会話か」である．図8にライブコンペ2において参加者が対話を評価している様子を示す．

＊5 オープントラックのNTTCSとは同組織であるが別チーム．

表 5　ライブコンペ 2 における各システムの手法．
順位に付与された T はタイを表す

| 順位 | チーム名 | 方式 | | | 知識源 | | | 学習手法 |
|---|---|---|---|---|---|---|---|---|
| | | ルール | 抽　出 | 生成 | テキスト | KB | 対　話 | |
| オープントラック | | | | | | | | |
| 1T | TokoChanTeam | ✓ | | | | ✓ | | N/A |
| 1T | UEC（Tripia） | ✓ | | | | | | N/A |
| 3 | NTTCS | ✓ | | | | ✓ | | SVM |
| 4T | UEC（NoneJupiter） | | ✓ | | ✓ | | | BERT |
| 4T | IRS | | ✓ | | | | ✓ | N/A |
| 6 | TEAM1 | ✓ | ✓ | | | | | N/A |
| 7 | KyoshiroS | | | ✓ | | | | RNN |
| 8T | MMI2019 | ✓ | ✓ | ✓ | | | ✓ | RNN |
| 8T | TEAM2 | | ✓ | ✓ | | | ✓ | RNN，SVM |
| シチュエーショントラック | | | | | | | | |
| 1 | OUHRI | ✓ | | | | | | CRF，ロジスティック回帰 |
| 2 | LiveCompetition2019 | ✓ | | | | | | N/A |
| 3 | NTTCS | ✓ | | | | ✓ | | Transformer |
| 4 | HSSLAB | ✓ | | | | | | N/A |
| 5 | 110 | ✓ | | | | | | N/A |
| 6 | nskw | ✓ | | | | | | N/A |

図 8　ライブコンペ 2 において参加者が対話を評価する様子

参加者による評価の結果，オープントラックでは 1 位は TokoChanTeam（3.61 点），2 位は NTTCS（3.28 点），3 位は UEC（Tripia）（2.81 点）であった．シチュエーショントラックは，1 位は OUHRI（3.70 点），2 位は NTTCS（3.48 点），3 位は LiveCompetition2019（3.23 点）であった．

図 9 は OUHRI のシステム（OHBOT）[中島 19] と対話者との対話である．絵文字が取り入れられており，人間らしい対話を実現しようとしている様子が垣間見える．ほとんどすべてのシステム発話が質問で終わっていることなど（これは，チャットボットを作成するときの常道であり，こうすることでユーザ発話を想定しやすくできる）人間らしくない点も見られるが，大きな破綻なく対話を続けることができている．

### 3·4　課　　題

対話システムの課題についてのアンケートの結果について述べる前に，ライブコンペ 2 では，システムのつくりや動作から明らかな問題が二つ見られた．

一つは，ライブコンペの予選ではクラウドワーカは一度しか対話をしないことを利用し，ルールベース，それも，システム主導によって一度きりの対話をうまく行うというシステムが予選を通過していたことである．これは，ライブコンペの予選を勝ち上がるためのハックのようなものである．具体的には，UEC（Tripia）[稲葉 19] がそうである．このシステムはユーザが何を発言しても，あらかじめ指定された 16 個の発話を順番に出力するというつくりのものであった．複数回システムと会話をするような予選であればよかったのかもしれないが，一度きりの対話を無難にこなすようにつくり込まれたシナリオによって予選が突破されてしまった*6．これは対話システムの評価の仕方について大きな課題を投げかけたといえる．

もう一つは，オープントラックとシチュエーショントラックの対話にあまり違いが見られなかったということである．具体的には，オープントラックのシステムであっても，最初からオープンに会話することを諦め，自身の得意なドメインに誘導するシステムが予選を勝ち残って

*6　シチュエーショントラックのオーガナイザによるベースライン（LiveCompetition2019）も同様のつくりをしていたが，こちらも予選を突破できた．

| 話者 | 発話 | ○ | △ | × |
|---|---|---|---|---|
| U | **/start** | | | |
| S | ところで，これまで行ったところで一番印象に残った場所ってどこ？ | 67.5% | 26.5% | 6.0% |
| U | **そうだねー、べただけどハワイかな** | | | |
| S | ハワイ！いつ行ったの？ | 91.6% | 8.4% | 0.0% |
| U | **うーん、もう結構前だよ** | | | |
| S | なるほど！旅先では、普段出会えないものや体験できないことが沢山あるよね😊ハワイはどうだった？ | 55.4% | 28.9% | 15.7% |
| U | **やっぱりきれいな海を見ながらのんびりできるのが最高だよね** | | | |
| S | そうなのね🙂ご当地グルメはなんか美味しいもの食べた？ | 71.1% | 21.7% | 7.2% |
| U | **そうだね、パシフィックリムっていうハワイの料理を食べたよ** | | | |
| S | うんうん！その旅行で一番印象に残ったのは？ | 28.9% | 49.4% | 21.7% |
| U | **うーん、海でのんびりしたのが一番印象的だったかなあ** | | | |
| S | そうなんだ！そういえば去年の夏沖縄に行ったよ。ゴーヤチャンプルがめちゃ美味しかった！ユウコもゴーヤ好き？ | 44.6% | 38.5% | 16.9% |
| U | **ゴーヤちょっと苦いけど好きだよ** | | | |
| S | そうなのね😐子供の時は食べれなかったけど、今はゴーヤのあのほろ苦さにハマってしまった😊😊 ビールなんかも最初は苦くて飲めなかったなあ。ユウコはどうだった？ | 72.3% | 19.3% | 8.4% |
| U | **そうだね、今でもビールよりはカクテルの方が飲みやすいかも** | | | |
| S | 沖縄で飲んだお酒よかったなあ。ちなみに沖縄の方言で「シマー」って何だかわかる？ | 38.6% | 34.9% | 26.5% |
| U | **えーわかんない** | | | |
| S | これね「泡盛」のことなんだよ！沖縄のことばってなんかほっこりするんだよね。「めんそーれー」とか「なんくるないさー」とか・・意味はよく分からないけど😉 | 81.9% | 13.3% | 4.8% |
| U | **あー、沖縄の方言って全然わかんないけど、ほっこりするね** | | | |
| S | あ、あとね、沖縄では締めにステーキ食べるんだって！ラーメンやらステーキやら、締め文化もいろいろだね！ユウコは知ってた？ | 55.4% | 31.3% | 13.3% |
| U | **えー、知らなかった。そんな重いのしめに食べるんだ。** | | | |
| S | うーん、こういう話してるとまた旅行に行きたくなる・・！ユウコは今行ってみたいところとかある？ | 51.8% | 32.5% | 15.7% |
| U | **うーん、またハワイに行ってみたいなあ** | | | |
| S | ほほう😊そのわけは？ | 78.3% | 14.5% | 7.2% |
| U | **うーん、今度は子供を連れて行ってみたい** | | | |
| S | 私ね、実は来年の春あたりに信州に行こうと思ってるんだ | 22.0% | 39.0% | 39.0% |
| U | **へーいいねー** | | | |
| S | 友達の実家が白馬で旅館やってて、遊びにおいでって誘ってもらってるんだ。よかったらユウコ一緒に行かない？ | 81.9% | 16.9% | 1.2% |
| U | **いきたーい** | | | |
| S | じゃあさ、今度ご飯でも食べながらゆっくり話そうよ😊 | 91.6% | 8.4% | 0.0% |
| U | **うん、そうだね、いいね！** | | | |
| S | いろいろ予定もあると思うから近いうちに決めようね😊 あ、時間になっちゃった😱話の途中なのにごめんね・・！じゃまた今度連絡するね | 86.7% | 13.3% | 0.0% |

図 9　対話者と OUHRI のシステム（OHBOT）との対話

いた．その結果，特定の話題の中だけで雑談を行うようなシステムばかりになってしまった．NTTCS は旅行ドメイン，TokoChanTeam は音楽ドメインであった．シチュエーショントラックのシステムについても，人間らしさを表出することは難しく，いわゆる「システム」的な応答を行うものが多かった．結果として，二つのトラックは似通ってしまったのである．これも大きくは評価の問題だといえるが，広い話題で破綻なく対話を続けること，人間らしい対話をモデリングすることの難しさを物語っているともいえる．

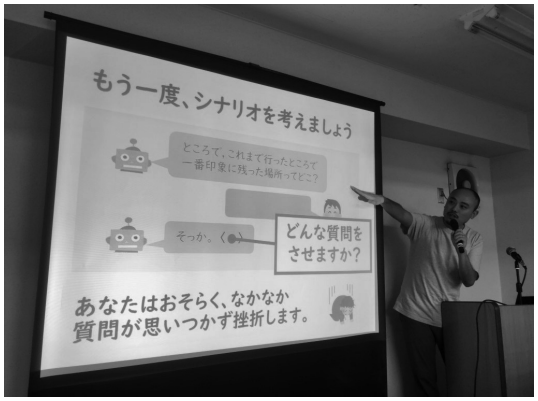



図 10　ライブコンペ講習会における飯尾氏のレクチャの様子

さて，ライブコンペ 1 と同様，ライブイベントの事前と事後に対話システムにおける課題についてのアンケートを実施した．アンケート項目は前回と同様（表 3）である．結果は表 6 のとおりである．1 位はライブコンペ 1 と同じく「発話内容の理解能力が低い／浅い」であったが，上述したような二つの問題を受けて，「広い話題に対応できない」や「品質評価が困難」が上位に位置している．また，シチュエーショントラックの対話を受けて，「ユーザに応じた応答ができていない」も順位を上げている．ライブコンペ 2 では，ライブコンペ 1 では見えてこなかった課題が浮き彫りにされたといえる．評価の仕方など仕様面での改善を図るとともに，これらの問題点を解決するシステムを今後のライブコンペでは期待したい．

表 6　ライブコンペ 2 後に参加者が重要と考えた対話システムの課題上位 10 件

| 順位 | 課題 | ランキングの変化 | |
|---|---|---|---|
| 1 | 【発話理解】発話内容の理解能力が低い／浅い | 1 => 1 | → |
| 2 | 【話題】広い話題に対応できない | 2 => 2 | → |
| 3 | 【ユーザ適応・個性】ユーザに応じた応答ができていない | 6 => 3 | ↗ |
| 4 | 【話題】適切な話題遷移ができない | 11 => 4 | ↗ |
| 5 | 【ビジネス・開発面】品質評価が困難 | 14 => 5 | ↗ |
| 6 | 【その他の機能】適切な主導権の交代ができない | 25 => 6 | ↗ |
| 7 | 【発話理解】質問に答える能力が低い | 12 => 7 | ↗ |
| 8 | 【発話生成】発話内容が文脈に沿っていない | 5 => 8 | ↘ |
| 9 | 【対話の意義】対話に面白さがない | 4 => 9 | ↘ |
| 10 | 【ユーザ適応・個性】過去のやり取りを活用できていない | 9 => 10 | ↘ |

## 4. 対話システムライブコンペ講習会

ここからはライブコンペに関連した二つのイベントについて述べる．一つは対話システムライブコンペ講習会である．

ライブコンペ 2 ではシチュエーショントラックを新たに導入したが，システムが人間らしい対話を行えるようにするにあたって，人文系の研究者の知見が有効であろうと考えた．例えば，会話分析，教育，言語学，社会学などの分野では人間どうしの対話を研究対象としている．これらの分野の研究者にライブコンペに参加してもらうことで工学系の研究者が思いもつかない豊かな対話を行うシステムが期待できると考えた．

しかし，人文系の研究者にとって，対話システムを一から構築することは難しいと考えられる．そこで，対話システムライブコンペ講習会と称して，人文系の研究者向けに対話システム構築のための講習会を行った．この講習会は，エントリの約 1 か月前，2019 年 9 月 8 日に実施した．30 名超の枠はすぐに埋まり，人文系の研究者に多く参加いただけた．

講習会の内容はルールベースによるシチュエーショントラック向け雑談対話システムの構築法である．対話ロボットのデモンストレーションを多く手掛けられている筑波大学の飯尾尊優氏に「対話システムにおけるシナリオ作成の秘訣」[*7] と題したレクチャを実施していただいた．内容としては，ユーザの発話にすべて応えようとするのはやめて，ユーザを質問責めにすることでユーザの発話を予測できるようにし，たいていの場合に適切に動作するルールを作成しようというものであった．これは，破綻のない対話システムの戦略としては王道であるが，前述のように，評価についての問題を引き起こすものでもある．

レクチャのあとは，Repl-AI [*8] と呼ばれる，ルールを GUI によって記述するツールの使い方とその演習，そして，最後に Repl-AI で作成したルールを Telegram から呼び出す方法についての講習を行った．

技術的な問合せが多くあったものの，本講習に参加した複数の人文系の研究者がライブコンペにエントリするなど一定の成果を実現することができた．また，参加者からは対話ルールの作成を日本語教育の題材に用いたいという声も聞かれた．対話ルールの作成は人文系と工学系をつなぐ良い接点となることがわかった．今後も講習会を開催し，人文系における対話の知見を対話システムに取り込んでいくとともに，対話システムの他分野への適用についても検討していきたい．

## 5. 日本語教育学会におけるパネルセッション

もう一つのイベントは日本語教育学会におけるパネルセッションである．このパネルは 2019 年度日本語教育学会秋季大会（2019 年 11 月 23 ～ 24 日開催）の初日に開催された．パネルを提案した目的は，人文系の研究者（今回は教育関係者）に対話システムの現状を知っていただき，工学系からは得られない人文系ならではの意見を聞きたいと考えたからである．

パネルではパネリストからの講演のほか，ライブコンペと同様の実演を行った．具体的には，ライブコンペ 1 において 1 位だった NTTCS のシステム（tripfreak）と 2 位だった NTTdocomo のシステム（MarikoZatsudanBot）[角森 18] と人間の話者との対話を参加者に鑑賞・評価してもらい，現状の対話システムの問題は何かについて討論を行った．加えて，ライブコンペ 1 と 2 でも実施した対話システムの課題についてのアンケートを実演の前後に実施した．アンケートには 60 ～ 80 名程度の方に記入いただけた．評価の結果は，ライブコンペ 1 の際と同様，NTTCS のシステムのスコア（2.80）が NTTdocomo のシステムのスコア（2.30）を上回った（スコアは高いほうが良い）．

表 7 に日本語教育学会の参加者が重要と考えた対話システムの課題上位 10 件を示す．アンケート結果は対話システムの研究者の結果と若干異なることが見て取れる．特に，「発話内容が文脈に沿っていない」，「適切な話題遷移ができない」が上位に入っている．また，「対話の目的が不明確」についても，実演の後，大きく順位が上昇している．人間どうしの対話に日頃から着目している教育関係者としては，文脈，話題，対話の目的（意図）といったものに適切に対応できていないシステムの挙動に違和感を覚えたのだろうと考えられる．

フリーフォームのアンケートでは，「反応が早すぎて気持ち悪い」，「質問の連続は飽きる」，「このままでは使いものにならない」，「会話ではない」などの意見が寄せられた．人文系の研究者がもつ対話についてのさまざまな知見を反映していくことで，より人間どうしの対話に近い対話が実現できるのではないかと考えており，今後もこのようなイベントを企画し，接点をもっていければと考えている．

**表 7** 日本語教育学会における実演後に参加者（主に教育関係者）が重要と考えた対話システムの課題上位 10 件

| 順位 | 課題 | ランキングの変化 | |
|---|---|---|---|
| 1 | 【発話生成】発話内容が文脈に沿っていない | 10 => 1 | ↗ |
| 2 | 【話題】適切な話題遷移ができない | 6 => 2 | ↗ |
| 3 | 【発話理解】発話内容の理解能力が低い／浅い | 5 => 3 | ↗ |
| 4 | 【ユーザ適応・個性】ユーザに応じた応答ができていない | 3 => 4 | ↘ |
| 5 | 【対話の意義】対話の目的が不明確 | 16 => 5 | ↗ |
| 6 | 【話題】広い話題に対応できない | 2 => 6 | ↘ |
| 7 | 【その他の機能】対話破綻から適切にリカバリできない | 4 => 7 | ↘ |
| 8 | 【発話理解】質問に答える能力が低い | 8 => 8 | → |
| 9 | 【発話理解】感情の理解能力が低い | 1 => 9 | ↘ |
| 10 | 【発話生成】日本語として不自然な文を生成する | 9 => 10 | ↘ |

＊7 https://youtu.be/GArldtAH8O0
＊8 https://repl-ai.jp/

## 6. ま と め

本稿では対話システムライブコンペティションについて紹介した．これまでに行った2回のライブコンペ（および，その関連イベント）について述べ，予選の結果や会場における参加者のアンケートの結果からわかった現状の対話システムの問題点について述べた．現在，特に問題だと考えられることは，システムの理解・応答能力の低さである．また，対話システムの評価の難しさも大きな課題である．

ライブコンペは，対話システムの関係者全員で対話システムの問題についての認識を共有でき，より良いシステム構築につなげることのできるイベントであると感じている．現在，ライブコンペ3に向けた準備をしているところであるが，次回はどのようなシステムが登場するのか，人文系の知見はどのようにシステムに生かされるのか，そして，どのような問題点が新たに浮かび上がるのか，今から楽しみである．

最後になるが，これまでの発表資料や講演動画，対話ログはホームページからダウンロードできるようにしている[*9, *10]．これらをぜひご覧いただき，興味をもっていただけたのであれば，ぜひライブコンペ3へのエントリを検討いただきたいと考えている．

### 謝 辞

ライブコンペの実施にあたっては，本学会より特別補助をいただきました．ライブコンペ講習会においては，飯尾尊優氏にご協力いただきました．日本語教育学会におけるパネルでは大塚容子氏にスムーズな進行をしていただきました．また，杉山弘晃氏にはtripfreakシステムをご準備いただきました．大変感謝いたします．

## ◇ 参 考 文 献 ◇

[Dinan 19] Dinan, E., Logacheva, V., Malykh, V., Miller, A., Shuster, K., Urbanek, J., Kiela, D., Szlam, A., Serban, I., Lowe, R., Prabhumoye, S., Black, A. W., Rudnicky, A., Williams, J., Pineau, J., Burtsev, M. and Weston, J.: The Second Conversational Intelligence Challenge（ConvAI2）, arXiv preprint, arXiv:1902.00098（2019）

[Henderson 14] Henderson, M., Thomson, B. and Williams, J. D.: The second dialog state tracking challenge, *Proc. SIGDIAL,* pp.263-272（2014）

[Higashinaka 16] Higashinaka, R., Funakoshi, K., Kobayashi, Y., andInaba, M.: The dialogue breakdown detection challenge: Task description, datasets, and evaluation metrics, *Proc. LREC,* pp. 3146-3150（2016）

[東中 18] 東中竜一郎，船越孝太郎，稲葉通将，角森唯子，高橋哲朗，赤間怜奈：対話システムライブコンペティション，SIG-SLUD, Vol. B5, No. 02, pp. 106-111（2018）

[東中 19a] 東中竜一郎：最近の対話システム事情：深層学習・データセット・コンペティションの観点から，映像情報メディア学会誌，Vol. 73, No. 2, pp. 271-276（2019）

[東中 19b] 東中竜一郎，船越孝太郎，稲葉通将，角森唯子，高橋哲朗，赤間怜奈，宇佐美まゆみ，川端良子，水上雅博：対話システムライブコンペティション2，第87回人工知能学会言語・音声理解と対話処理研究会（第10回対話システムシンポジウム）（2019）

[Higashinaka 19c] Higashinaka, R., Funakoshi, K., Inaba, M., Tsunomori, Y., Takahashi, T. and Akama, R.: Dialogue system live competition: Identifying problems with dialogue systems through live event, *Proc. IWSDS*（2019）

[東中 20] 東中竜一郎，稲葉通将，水上雅博：Pythonでつくる対話システム，オーム社（2020）

[Hori 19] Hori, C., Perez, J., Higashinaka, R., Hori, T., Boureau, Y.-L., Inaba, M., Tsunomori, Y., Takahashi, T., Yoshino, K. and Kim, S.: Overview of the sixth dialog system technology challenge: DSTC6, *Computer Speech & Language,* Vol. 55, pp. 1-25（2019）

[稲葉 19] 稲葉通将：雑談対話システムをどう評価すべきか―TripiaBotのライブコンペ予選通過から考える―，第87回人工知能学会言語・音声理解と対話処理研究会（第10回対話システムシンポジウム）（2019）

[角森 18] 角森唯子，大西可奈子，藤本拓，角野公亮，吉村 健，礒田佳徳：カスタマイズ可能なオープンドメイン雑談対話エンジンの開発，第84回人工知能学会言語・音声理解と対話処理研究会（第9回対話システムシンポジウム）（2018）

[中島 19] 中島圭祐，駒谷和範，中野幹生：雑談対話システム構築フレームワークPyChatに基づく特定シチュエーション向け対話システム，第87回人工知能学会言語・音声理解と対話処理研究会（第10回対話システムシンポジウム）（2019）

[中野 15] 中野幹生，駒谷和範，船越孝太郎，中野有紀子：対話システム，コロナ社（2015）

[杉山 18] 杉山弘晃，成松宏美，水上雅博，有本庸浩：文脈に沿った発話理解・生成を行うドメイン特化型雑談対話システムの実験的検討，第84回人工知能学会言語・音声理解と対話処理研究会（第9回対話システムシンポジウム）（2018）

[宇佐美 19] 宇佐美まゆみ，東中竜一郎，杉山弘晃，角森唯子，高橋哲朗，大塚容子：対話システム構築と談話研究・日本語教育の接点，2019年度日本語教育学会秋季大会予稿集（2019）

2020年3月2日 受理

---

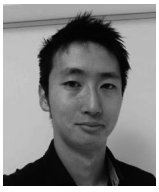
＊9 https://dialog-system-live-competition.github.io/dslc1/index.html

＊10 https://dialog-system-live-competition.github.io/dslc2/index.html

## 著 者 紹 介

**東中 竜一郎**（正会員）

2001 年慶應義塾大学大学院政策・メディア研究科修士課程，2008 年博士課程修了．2001 年日本電信電話株式会社入社．現在，NTT メディアインテリジェンス研究所上席特別研究員．質問応答システム・音声対話システムの研究開発に従事．博士（学術）．「しゃべってコンシェル」の質問応答機能の研究開発や，「ロボットは東大に入れるか」プロジェクトにおける英語科目を担当．平成 28 年度科学技術分野の文部科学大臣表彰を受賞．著書に「質問応答システム」（コロナ社，2009），「おうちで学べる人工知能のきほん」（翔泳社，2017），「人工知能プロジェクト「ロボットは東大に入れるか」：第三次 AI ブームの到達点と限界」（東京大学出版会，2018），「Python でつくる対話システム」（オーム社，2020）など．言語処理学会，情報処理学会，電子情報通信学会各会員．

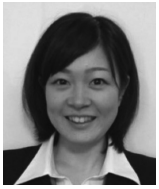


**船越 孝太郎**（正会員）

2000 年東京工業大学工学部情報工学科卒業．2002 年同大学院情報理工学研究科計算工学専攻修士課程修了．2005 年同研究科博士課程修了．同年，東京工業大学大学院特別研究員．2006 年（株）ホンダ・リサーチ・インスティチュート・ジャパン入社．2013 年より同シニア・リサーチャ．2017 年より京都大学大学院情報学研究科知能情報学専攻特定准教授（出向）．博士（工学）．自然言語理解，マルチモーダル対話に関する研究に従事．情報処理学会，言語処理学会，ACM SIGCHI 各会員．

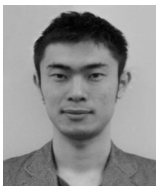
**稲葉 通将**（正会員）

2012 年名古屋大学大学院情報科学研究科博士後期課程修了．同年，広島市立大学大学院情報科学研究科助教，2019 年電気通信大学人工知能先端研究センター准教授．現在に至る．博士（情報科学）．対話システム，対話処理に関する研究に従事．電子情報通信学会，情報処理学会，言語処理学会各会員．



**角森 唯子**（正会員）

2015 年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了．同年，（株）NTT ドコモ入社．自然言語処理および対話システムに関する研究開発に従事．



**高橋 哲朗**

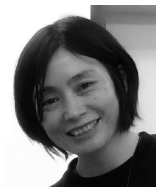
2005 年奈良先端科学技術大学院大学博士後期課程修了．博士（工学）．同年，（株）富士通研究所入社．2008 ～ 10 年ニフティ株式会社にて Web サービス開発，2011 ～ 12 年および 2014 ～ 15 年マサチューセッツ工科大学にて客員研究員を経て，現在，（株）富士通研究所にて自然言語処理およびデータ分析の研究に従事．言語処理学会，情報処理学会各会員．

**赤間 怜奈**（学生会員）

2017 年東北大学工学部卒業．2018 年同大学院情報科学研究科博士前期課程修了．現在，同研究科博士後期課程在学中．自然言語処理に関する研究に従事．言語処理学会，ACL 各学生会員．

**宇佐美 まゆみ**

東京外国語大学大学院教授などを経て，2016 年度より国立国語研究所日本語教育研究領域教授．教育学博士，Ed.D（ハーバード大学）．専門は，言語社会心理学，日本語教育学．慶應義塾大学大学院社会学研究科心理学専攻，ハーバード大学教育学部大学院言語・文化修得専攻博士課程修了．ディスコース・ポライトネス理論研究，自然会話分析のほか，現在は，その知見を対話システム研究に生かせないか模索中．主な著書・論文に，「言葉は社会を変えられる」（編著）（明石書店，1997），「Discourse politeness in Japanese conversation: Some implications for a universal theory of politeness」（Hituzi Syobo, 2002）．「ポライトネス理論の展開（1 ～ 5，7 ～ 13）」『月刊言語』（1 ～ 12 月号）（大修館書店，2002）など．「ポライトネス理論研究のフロンティア－ポライトネス理論研究の課題とディスコース・ポライトネス理論」『社会言語科学』，Vol. 11, No. 1, pp. 4-22（2008），「相互作用と学習―ディスコース・ポライトネス理論の観点から」『西原鈴子･西郡仁朗 編：講座社会言語科学』, 4 巻, 教育･学習（ひつじ書房, 2008），「談話のポライトネス－ポライトネスの談話理論構想－」『談話のポライトネス』（第 7 回国立国語研究所国際シンポジウム報告書）（国立国語研究所，2001），などがある．



**川端 良子**（正会員）

国立国語研究所音声言語研究領域プロジェクト非常勤研究員．2002 年千葉大学大学院文学研究科修了．2019 年同大学院融合科学研究科博士課程修了．博士（学術）．2016 年から現職．「大規模日常会話コーパスに基づく話し言葉の多角的研究」に参加し，コーパスの開発に取り組んでいる．対話において共有信念がどのような仕組みで更新されるかに興味をもっており，コーパスを用いた研究を行っている．

**水上 雅博**

2014 年奈良先端科学技術大学院大学情報科学研究科修士課程，2017 年博士課程修了．2017 年日本電信電話株式会社入社．現在，NTT コミュニケーション科学基礎研究所研究員．雑談対話システムの研究開発に従事．博士（工学）．