

国際会議報告

第11回対話システムシンポジウム (2020/12/1)

報告する会議

- SIGDIAL 有本 庸浩 (NTTコミュニケーション科学基礎研究所)
- ACL 佐藤 志貴 (東北大学)
- ICMI 井上 昂治 (京都大学)
- INTERSPEECH 山本 賢太 (京都大学)
- EMNLP 赤間 恽奈 (東北大学)

第11回対話システムシンポジウム
2020年12月1日



SIGDIAL 2020 国際会議報告

有本 庸浩
(NTT)

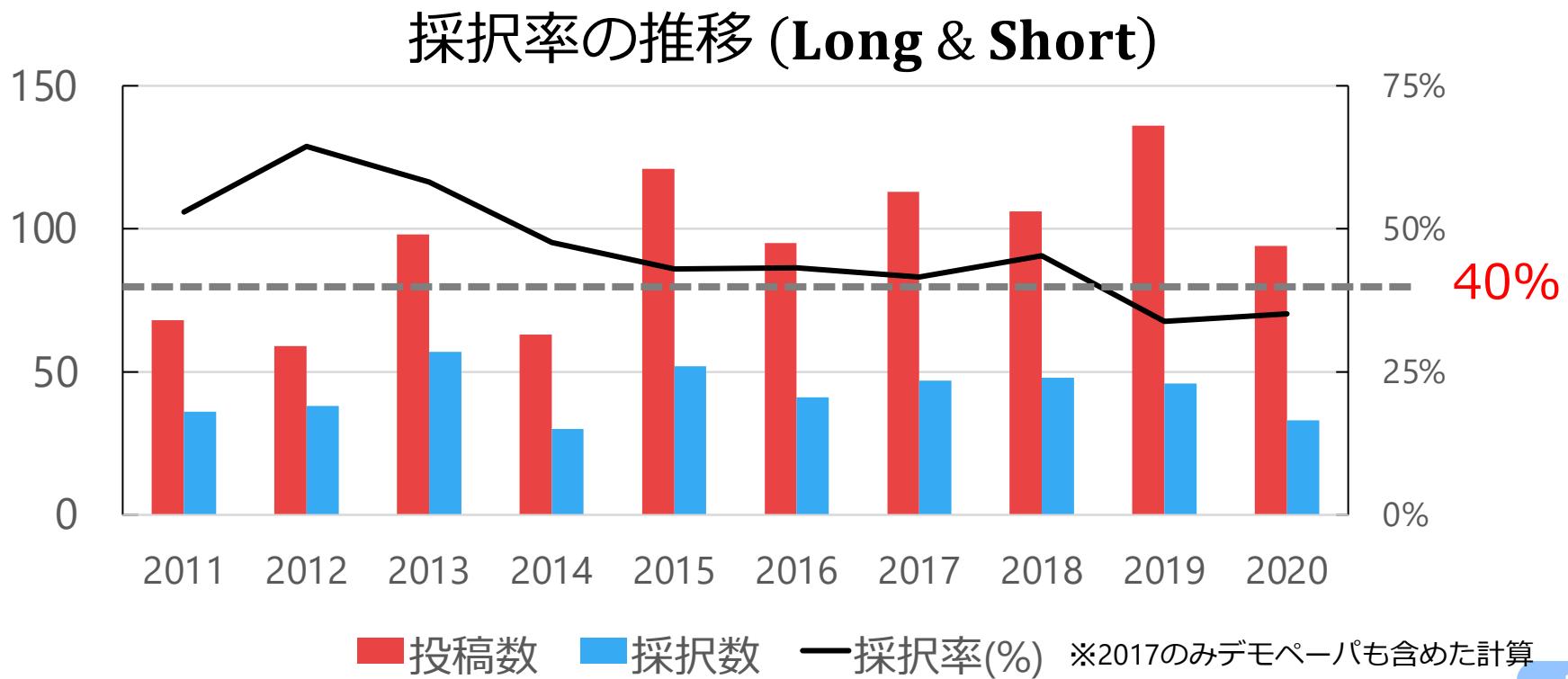
SIGDIAL とは

(Special Interest Group on Discourse and Dialogue)

- 対話システムや対話分析に関する発表が集まる会議
 - 談話処理
 - 対話システム
 - コーパス収集・分析
- SIGDIAL 2020
 - 2020年7月1日～3日
 - バーチャル
 - 参加登録人数 980人, youtube再生回数 5.4K +

採択率 (Long & Short): 35%

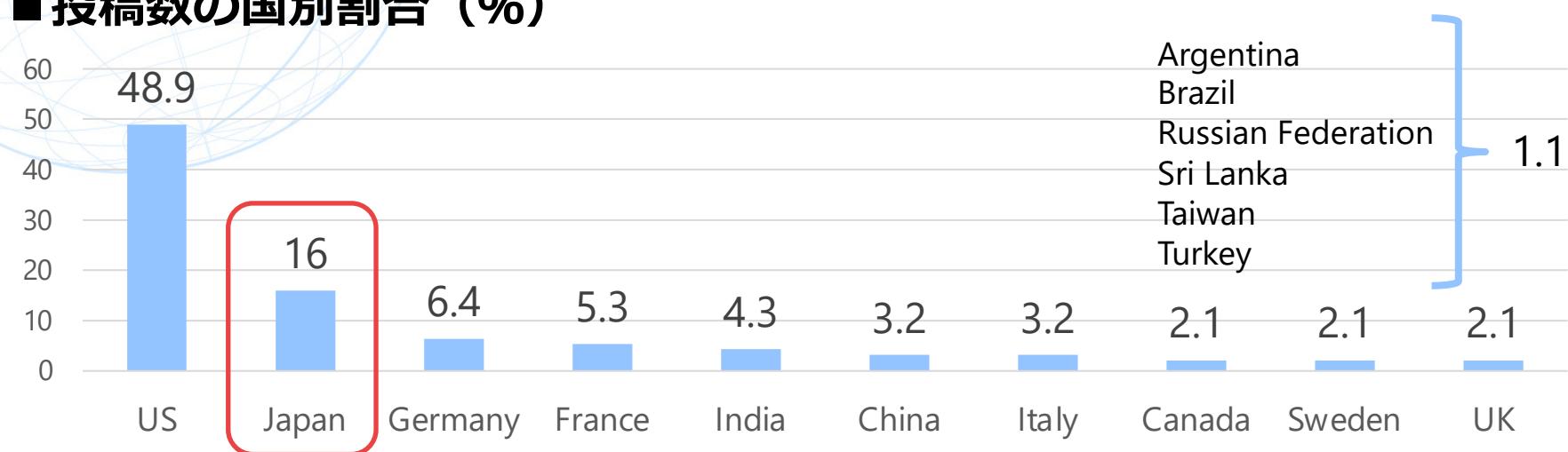
- 昨年度と同程度の難易度
- 投稿種別
 - Long: 37% (23/62)
 - Short: 31% (10/32)
 - Demo: 80% (8/10)



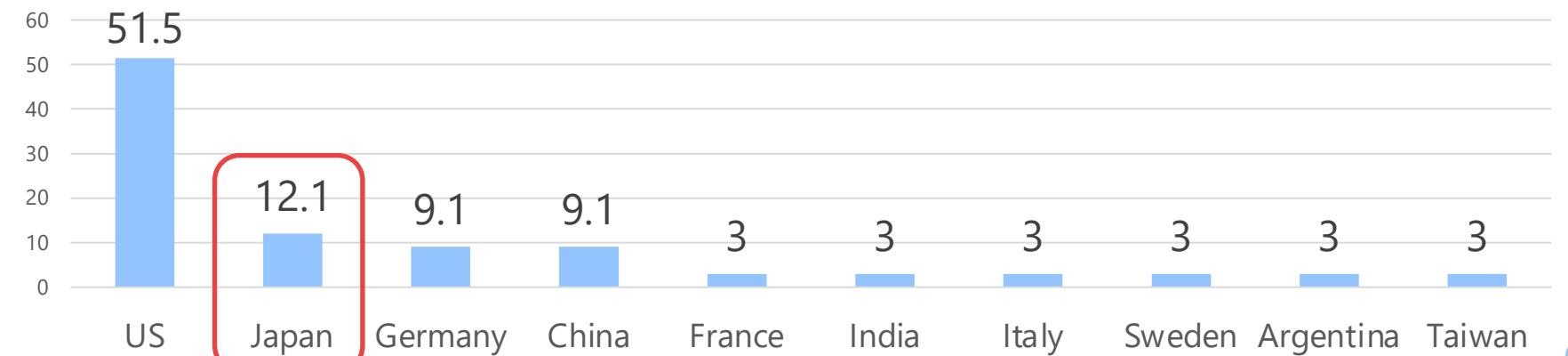
国別の参加動向

- アメリカの投稿が約半分。日本は2番目に多い。

■ 投稿数の国別割合 (%)



■ 採択数の国別の割合 (%)



バーチャル開催の様子 (1/2)

Trello (メイン会場)

The screenshot shows the SIGDIAL 2020 Trello board. The left sidebar contains sections for Conference Management, Schedule (with three days of streams), and General Information. The main board area is divided into several columns representing different types of sessions:

- Keynotes:** Includes sessions like "Keynote 1: Gabriel Skantze" (20 views, 15 comments) and "Keynote 2: Adil Celikyilmaz" (19 views, 1 comment).
- Session 1: Generation & Task-oriented Dialogue 1:** Includes "Semantic Guidance of Dialogue Generation with Reinforcement Learning" by Cheng-Hsun Hsueh and Wei-Yun Ma.
- Session 2: Demo 1:** Includes "Conversational Agents for Intelligent Buildings" by Weronika Sieniak, Christian Dondrup, Nancie Gunson, and Oliver Lemon.
- Session 3: Knowledge Use and Behaviour 1:** Includes "RETICO: An incremental framework for spoken dialogue systems" by Michael Michael.
- Discourse Processing 2 (Carolyn Rose):** Includes "Learning from Mistakes: Combining Ontologies via Self-Training for Dialogue Generation" by Lena Reed, Vrindavan Harrison, Shereen Oraby, Dilek Hakkani-Tur, and Marilyn Walker.
- Chatbots (Ryuichiro Higashinaka):** Includes "TripPy: A Triple Copy Strategy for Value Independent Neural Dialog State Tracking" by Michael Heck, Carel van Niekerk, Nurul Lubis, Christian Geisbauer, Hsien-Chin Lin, Marco Moresi, and Milica Gasic.
- Best practices in human evaluation of dialogue and NLG - More art than science? (Verena Rieser):** Includes "Convolt: A Toolkit for the Analysis of Conversations" by Jonathan P. Chang, Caleb Chiam, Liye Fu, Andrew Wang, Justine Zhang, and Cristian Danescu-Niculescu-Mizil.

Youtube
+ Zoom

動画視聴

ライブQA

Discussion

- 動画時間: ロングが12分, ショートが8分, デモが10分

※ 公開されました

<https://trello.com/b/tFM0R8G3/sigdial-2020>

https://www.youtube.com/channel/UCpT0_wXw7cScSnElXQQTxYg/playlists

バーチャル開催の様子(2/2)

- 研究者らによるブレイクアウトディスカッション(各1時間程度)

- What Newcomers Should Know about Dialogue
- Personalization
- Turn-taking
- Natural language Generation (NLG)
- Discourse Processing
- Chatbots
- Best practices in Human evaluation of dialogue and NLG
- Dialogue Management
- Dialogue with Robots
- Natural Language Understanding
- Multimodal Dialogue
- Evaluation
- Multi-party Dialogue
- Spoken Language Understanding
- Discourse Processing
- REAL Dialogue Challenge
- Virtual Conversational Agents in the Text Channel
- Emotion and Impression in Dialogue
- Human-human Dialogue Models

Keynote

- “Conversational Turn-taking in Human–robot Interaction”
Gabriel Skantze (KTH Royal Institute of Technology)
円滑なターンテーキングのための言語、韻律、視線情報の活用について：
現状のロボットは、ターンテーキングを同期する情報受信・生成が課題。
- “Neural Text Generation: Progress and Challenges”
Asli Celikyilmaz (Microsoft Research)
ニューラルベースのテキスト生成の既存技術について：
制御(会話スタイル、感情)、知識活用、常識、対話モデリング
- “Argument Mining, Discourse Analysis, and Educational Applications” Diane Litman (University of Pittsburgh)
議論構造のマイニング・教育のための活用について

セッション

Generation & Task-oriented Dialogue	8
Knowledge Acquisition, Use and Behavior Modeling	6
Evaluation & Corpora	6
Semantic Pragmatics Modeling	4
Situated Dialogue with Virtual Agents and Robots (Special Session)	4
Modeling Convergence	3
Dialogue Policy	2
Demo	8
合計	41

- 対話システムに関する話題は出しやすい印象。性能だけではなく、対話の本質的な課題に対する妥当な解決策・分析をちゃんと示せたかが重要な印象。

発表の紹介 (1/2)

タスク指向型対話

- 対話状態追跡のスロットフィリングは、未知の値に弱い。発話から値を抽出・コピーする手法で改善
 - TripPy: A Triple Copy Strategy for Value Independent Neural Dialog State Tracking (ベストペーパー)
- タスク指向型対話システムは、特定の設定・モジュールに対する評価のみになりやすい。システム全体の構成や設定、評価方法が与える影響を分析
 - Is Your Goal-Oriented Dialog Model Performing Really Well? Empirical Analysis of System-wise Evaluation (ベストペーパー)
- NLGにおいて、新しいドメインのオントロジーに対応することはコストがかかる。既知オントロジーを組み合わせるための自己訓練手法を提案
 - Learning from Mistakes: Combining Ontologies via Self-Training for Dialogue Generation

発表の紹介 (2/2)

知識獲得・利用

- 語彙知識を得るためのシステム質問の印象評価
 - User Impressions of Questions to Acquire Lexical Knowledge
- 常識に基づく読解
 - Commonsense Evidence Generation and Injection in Reading Comprehension

振る舞いのモデリング

- カスタマーサービスの人間スタッフ/ボットに対するユーザの振る舞い差を分析
 - A Case Study of User Communication Styles with Customer Service Agents versus Intelligent Virtual Agents

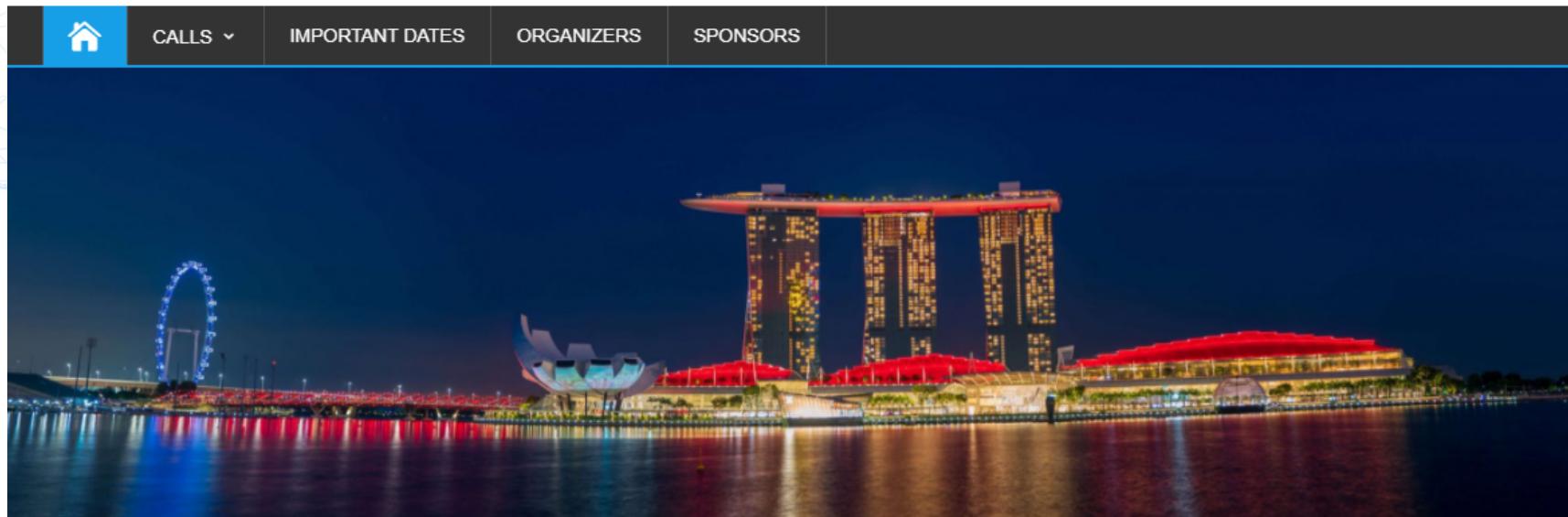
対話評価

- DialoGPTを用いた、対話の自動評価尺度
 - Unsupervised Evaluation of Interactive Dialog with DialoGPT

来年度の情報

SIGDIAL 2021

The 22nd Annual Meeting of the Special Interest Group on Discourse and Dialogue, 29-31 July 2021, Singapore



<https://www.sigdial.org/files/workshops/conference22/>

- 開催期間: 2021.07.29–31
- 開催場所: シンガポール
- 投稿締切は3月頃？

国際会議報告 ACL2020

東北大大学 情報科学研究科 修士2年

佐藤志貴

shiki.sato@ecei.tohoku.ac.jp

2020/12/01

The 58th Annual Meeting of the Association for Computational Linguistics



Design by Jingya Chen

ACL2020 @online, 7/5-7/8

- 計算言語学・自然言語処理の国際会議のひとつ
 - Association for Computational Linguisticsにより年一回開催
- 計算科学を用いて自然言語を扱う幅広い研究が発表

今日お話しする内容

- オンライン開催について
- 論文の統計と傾向
- 個人的に注目したトピック

オンライン開催での発表形式

発表者がやったことは以下の三つ

1. 発表ビデオの録画

- ・ 口頭発表の様子を事前に録画
- ・ **long: 12分, short: 7分** (ACL2019はlong: 15分, short: 10分)

2. ライブQ&Aセッションでの質疑応答対応

- ・ 発表者ごとのZoom部屋で待機し質問に対応
- ・ 会議期間中の1時間×2スロット

3. テキストチャットでの質疑応答対応

- ・ Rocket.Chatに来る質問に返信
- ・ 会議期間中いつでも質問&いつでも返信

オンライン発表を終えての反省等

- ビデオ撮影について

- 撮影場所の確保、撮影、投稿ともに想定以上の時間がかかった。
撮影や投稿に使うサービスは学会ごとに異なり手間取る可能性
- ビデオの字幕提出依頼に備え正確な原稿を用意しておくと便利
(撮影時に原稿を使わない方も含め)

- ライブQ&Aセッションについて

- (事前準備) オンラインなら質問がどうしても聞き取れない時は
チャット欄も使えるので「聞く」より「話す」練習を優先
- 30秒程の研究概要を用意しておくと、色々なセッションを見て
回っているビデオ未視聴の人との議論のきっかけに

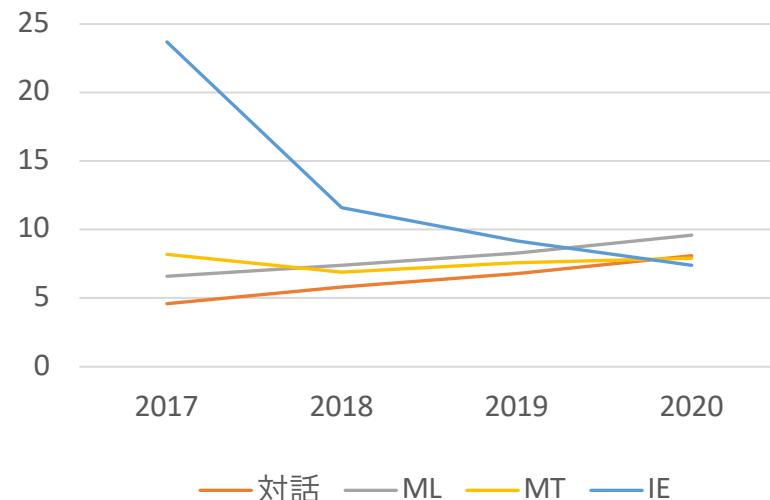
論文の投稿・採択に関する統計

投稿数および採択数 (カッコ内は対話エリアでの数字)

	Total Submissions	Accepted	% Accepted
ACL 2017	1297 (60)	302 (16)	23.3 (26.7)
ACL 2018	1544 (90)	384 (23)	24.9 (25.6)
ACL 2019	2680 (183)	660 (52)	24.6 (28.4)
ACL 2020	3088 (250)	779 (62)	25.2 (24.8)

※査読された論文のみを集計

各エリアの投稿数割合



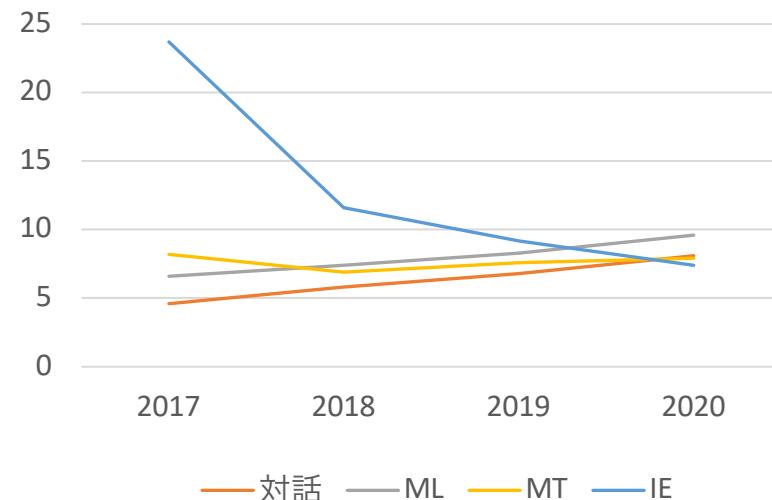
論文の投稿・採択に関する統計

投稿数および採択数 (カッコ内は対話エリアでの数字)

	Total Submissions	Accepted	% Accepted
ACL 2017	1297 (60)	302 (16)	23.3 (26.7)
ACL 2018	1544 (90)	384 (23)	24.9 (25.6)
ACL 2019	2680 (183)	660 (52)	24.6 (28.4)
ACL 2020	3088 (250)	779 (62)	25.2 (24.8)

※査読された論文のみを集計

各エリアの投稿数割合



- 全体の投稿数は**増加傾向**

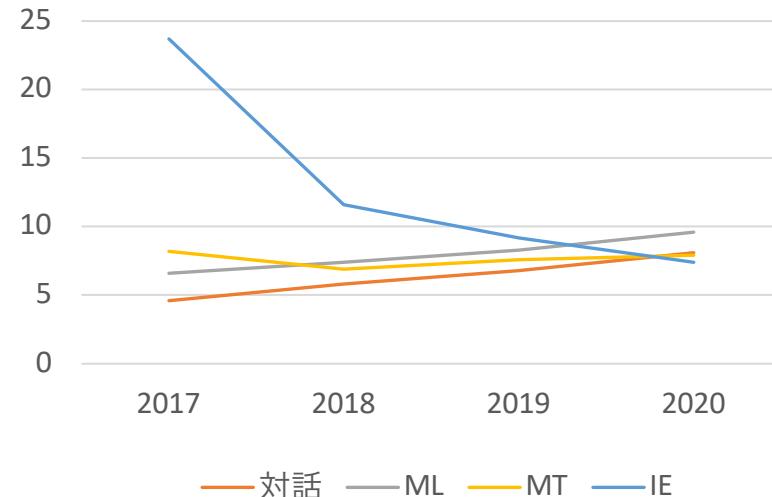
論文の投稿・採択に関する統計

投稿数および採択数 (カッコ内は対話エリアでの数字)

	Total Submissions	Accepted	% Accepted
ACL 2017	1297 (60)	302 (16)	23.3 (26.7)
ACL 2018	1544 (90)	384 (23)	24.9 (25.6)
ACL 2019	2680 (183)	660 (52)	24.6 (28.4)
ACL 2020	3088 (250)	779 (62)	25.2 (24.8)

※査読された論文のみを集計

各エリアの投稿数割合



- 全体の投稿数は増加傾向
- 対話の投稿数も**増加傾向**

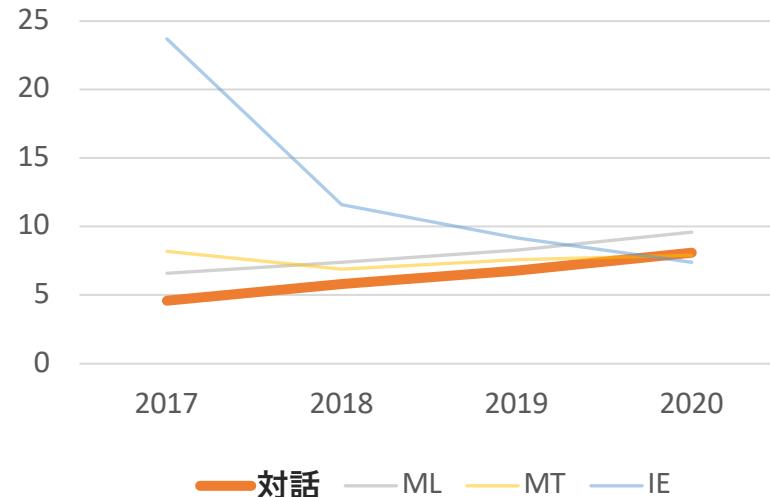
論文の投稿・採択に関する統計

投稿数および採択数 (カッコ内は対話エリアでの数字)

	Total Submissions	Accepted	% Accepted
ACL 2017	1297 (60)	302 (16)	23.3 (26.7)
ACL 2018	1544 (90)	384 (23)	24.9 (25.6)
ACL 2019	2680 (183)	660 (52)	24.6 (28.4)
ACL 2020	3088 (250)	779 (62)	25.2 (24.8)

※査読された論文のみを集計

各エリアの投稿数割合



- 全体の投稿数は増加傾向
- 対話の投稿数も増加傾向
- 全体に占める対話論文の割合も上昇傾向

採択論文の傾向

Dialogue and Interactive Systemsで発表された論文61本を題目と概要をもとに分類

- **応答生成: 27**

- 学習方法: 12
- アーキテクチャ改良: 7
- 応答への知識付与: 6
- マルチモーダル: 2

- **対話管理: 12**

- **対話理解: 8**

- **評価: 8**

- **その他: 6**

採択論文の傾向

Dialogue and Interactive Systemsで発表された論文61本を題目と概要をもとに分類

- 応答生成: 27

- 学習方法: 12
- アーキテクチャ改良: 7
- 応答への知識付与: 6
- マルチモーダル: 2

昨年同様ニューラル応答生成に
関する論文が支配的

なかでも学習方法の提案に
関する論文が多数

- 対話管理: 12

- 対話理解: 8

- 評価: 8

- その他: 6

採択論文の傾向

Dialogue and Interactive Systemsで発表された論文61本を題目と概要をもとに分類

• 応答生成: 27

- 学習方法: 12
- アーキテクチャ改良: 7
- 応答への知識付与: 6
- マルチモーダル: 2

昨年同様ニューラル応答生成に
関する論文が支配的

なかでも学習方法の提案に
関する論文が多数

• 対話管理: 12

• 対話理解: 8

• 評価: 8

• その他: 6

システムやその応答の評価に
関する論文が多数

採択論文の傾向

Dialogue and Interactive Systemsで発表された論文61本を題目と概要をもとに分類

応答生成: 27

- 学習方法: 12
- アーキテクチャ改良: 7
- 応答への知識付与: 6
- マルチモーダル: 2

昨年同様ニューラル応答生成に関する論文が支配的

対話管理: 12

対話理解: 8

評価: 8

その他: 6

なかでも学習方法の提案に関する論文が多数

二つのトピックに注目

システムやその応答の評価に関する論文が多数

注目トピック1 応答生成/学習方法

- ・多数のタスクやデータが対話研究のなかで提案されてきた
- ・それらを応答生成モデルの学習にどう活用するかに関する論文が多数採択

応答生成:

- ・学習方法: 12
 - ・マルチタスク: 4
 - ・学習データの扱いの工夫: 3
 - ・Negative/Unlikelihood学習: 2
 - ・**マルチスキル: 2** ←————
 - ・メタ学習: 1

今回は特に
このトピックに注目

注目トピック1 応答生成/学習方法

マルチスキル

知識付与・ペルソナ付与など、個別に研究されてきた複数の応答生成の「スキル」を備えたシステムを学習

- Shuster+'20 The Dialogue Dodecathlon: Open-Domain Knowledge and Image Grounded Conversational Agents
 - 12個のデータセットに関するスキルの集約について検証
 - マルチタスク学習の有効性を確認
- Smith+'20 Can You Put it All Together: Evaluating Conversational Agents' Ability to Blend Skills
 - 3個のデータセットに関するスキルの集約について検証
 - マルチタスク学習の有効性を確認

感想等

- 近年は特色のある小規模 but 高品質対話データが多数登場
- これらの活用方法に関する研究は今後も盛り上がる可能性

注目トピック2 評価

- 評価系論文8本中4本はシステム応答の自動評価
- DNNベースのシステム応答自動評価が多数採択
 - [Mehri+'20 USR: An Unsupervised and Reference Free Evaluation Metric for Dialog Generation](#)
 - [Zhao+'20 Designing Precise and Robust Dialogue Response Evaluators](#)
 - [Sinha+'20 Learning an Unreferenced Metric for Online Dialogue Evaluation](#)
 - [Bak&Oh'20 Speaker Sensitive Response Evaluation Model](#)

注目トピック2 評価

DNNによる応答自動評価

- [Zhao+'20 Designing Precise and Robust Dialogue Response Evaluators](#)
 - DNNモデルベース応答自動評価の精度向上方法の提案と検証
 - RoBERTa[Liu+'16]ベースの評価法が人手評価と強く相関することを確認

Model	Pr.	Spr.	SD
Automated Metrics			
BLEU-2	0.31	0.23	0.31
Average	0.25	0.23	0.19
Extrema	0.26	0.26	0.23
Greedy	0.25	0.23	0.21
Baseline Evaluator			
ADEM	0.34**	0.36**	0.51
RUBER	0.37**	0.31*	0.67
Proposed Evaluator			
RUBER			
<i>semi-sup.</i>	0.45**	0.41**	0.42
<i>unref.+semi-sup.</i>	0.43**	0.39**	0.83
RoBERTa-eval	0.64**	0.66**	1.26
Human Judgement			
Human	1.0	1.0	1.42

Table 3: Performances of automated metrics, baseline evaluators, and proposed evaluators (§6).

感想等

- 巨大LMを使っても人手評価にかわる評価は困難
- 自動評価では捉えられない側面を明らかにした上で人手評価との共存など有効な利用法を模索することが重要

**The Joint Conference of the 59th Annual Meeting of the Association
for Computational Linguistics and the 11th International Joint
Conference on Natural Language Processing (ACL-IJCNLP 2021)**



Photo by Milkovi on Unsplash

ACL2021 @Bangkok, 8/1-8/6

- Submission Deadline: **2021/02/01**
- Author Response : 2021/03/25-31
- Notification : 2021/05/05



International Conference on Multimodal Interaction 2020 (ICMI 2020)

第11回対話システムシンポジウム

国際会議報告

井上 昂治 (京都大学)

2020年12月1日

<http://www.sap.ist.i.kyoto-u.ac.jp/members/inoue/>

ICMI とは

- マルチモーダルインタラクションに関するACM系の国際会議
 - “the premier international forum for multidisciplinary research on multimodal human-human and human-computer interaction, interfaces, and system development”
 - “**Scientific Novelty** と **Technical Novelty** の両方を重視”
- 会議の規模：毎回100件弱の発表 (“密”なインタラクション)
- この会議も今年は例によってオンライン開催へ…

ICMI 2020 の統計

- 投稿数 **223** (long 159, short 64)
- 採択数 **84 (38%)** (long 65 = 41%, short 19=30%)

- Keynote (**4**)
- Tutorial (**1**)
 - "Conversational interaction with social robots" by Furhat Robotics
- Grand challenge (**1**)
 - Eighth Emotion Recognition in the Wild Challenge (4 subtopics)
- Demonstrations/Exhibit (**6**)
- Doctoral Consortium (**12**)
- Late-Breaking Results paper (**23**)
- Workshop (**13**)

会議の開催方法

- Ryver と Zoomを利用
 - Keynote と Best paper nominate は Zoomのメイン会場でオーラル形式
 - その他の発表は Zoom の ブレークアウトルームでポスター形式

Time CET/ UTC+1	Workshop Day 25 Oct	Day 1 26 Oct	Day 2 27 Oct	Day 3 28 Oct	Workshop Day 29 Oct
13:15					
14:15					
14:15 - 15:00	Workshops: DVU2020 FGAHI IGTD20 MHFI Mi-Psy'20 Doctoral Consortium & Grand Challenge	Opening & Virtual Tour Utrecht	Breakout Discussions #1	Drinks Pub Quiz	Workshops: WoCBU MMEC MSMT2020 & Furhat Tutorial
15:00 - 16:00		Keynote Asli Ozyurek	Best Paper Nominees #2	Keynote Atau Tanaka	
16:00 - 16:45		Session 1a Papers, DC & Demos	Session 2a Papers & LBR	Session 3a Papers & GC	
16:45-17:00	Break	Break	Break	Break	Break
17:00 - 17:45		Session 1b Papers, DC & Demos	Session 2b Papers & LBR	Session 3b Papers & GC	Workshops: MAStroPE MSECPE2020 OHT SAMIH & Furhat Tutorial
17:45 - 18:45		Best Paper Nominees #1	Keynote Shrikanth Narayanan	Keynote Kate Saenko	
18:45 - 19:30	Doctoral Consortium & Grand Challenge	Open- stream Drinks	Pub Quiz	Breakout Discussions #2	Town Hall Meeting, Awards & Closing
19:30 - 20:15					

The screenshot shows the Ryver platform interface. On the left, there is a sidebar with the user's profile picture, name (@kojinoue), and a dropdown menu. Below the profile are sections for Favorites, Schedule, Forums, Teams, and People. The main area is titled "Schedule" and contains a grid of tasks. Each task card includes a timestamp, title, description, and a "Zoom" button. The tasks are categorized into sessions and keynotes. Some tasks havehtags like #Day1, #Zoom, #Short, or #Long.

Session / Keynote	Date	Description	Hashtags
Session 1a Make up: Long, Short, DC, Demo	Thu 29 Oct 16:00-16:45	(CET/UTC+1) Zoom link	#Zoom #make-up #Day1 #a month ago
Session 1b Make up: Long, Short, DC, Demo	Thu 29 Oct 17:00-17:45	(CET/UTC+1) Zoom link	#Zoom #make-up #Day1 #a month ago
Keynotes	Mon 26 Oct 15:00-16:00	(CET/UTC+1) Keynote 1: Asli Ozyurek	#KEY1 #Zoom #Day1 #a month ago
Keynotes	Mon 26 Oct 16:00-16:45	(CET/UTC+1) Keynote 2: Shrikanth Narayanan	#KEY2 #Zoom #Day1 #a month ago
Keynotes	Tue 27 Oct 17:45-18:45	(CET/UTC+1) Keynote 3: Atau Tanaka	#KEY3 #Zoom #Day2 #a month ago
Keynotes	Wed 28 Oct 15:00-16:00	(CET/UTC+1) Keynote 4: Kate Saenko	#KEY4 #Zoom #Day3 #a month ago
Session 3a: Long, Short & Grand Challenge papers	Wed 28 Oct 16:00-16:45	(CET/UTC+1) Zoom link	#S3A1 #Zoom #Day3 #a month ago
Session 3b: Long, Short & Grand Challenge papers	Wed 28 Oct 17:00-17:45	(CET/UTC+1) Zoom link	#S3B1 #Zoom #Day3 #a month ago
Session 3a: Long, Short & Grand Challenge papers	Wed 28th @ 17:00-17:45	Daily Furhat Demo: Wed 28th @ 16:00-16:45	#S3A2 #Zoom #Demo #Exhibit #Day2 #a month ago
Session 3b: Long, Short & Grand Challenge papers	Wed 28th @ 17:00-17:45	Daily Furhat Drop-in Session: Wed 28th @ 17:00-17:45	#S3B2 #Zoom #Demo #Day3 #a month ago
Session 3a: Long, Short & Grand Challenge papers	New Task...		
Session 3b: Long, Short & Grand Challenge papers	New Task...		

Keynote

- “From Hands to Brains: How Does Human Body Talk, Think and Interact in Face-to-Face Language Use?”
Asli Özyürek (MaxPlanck Institute)
- “Deep Learning for Joint Vision and Language Understanding”
Keito Saenko (Boston University)
- “Musical Multimodal Interaction: From Bodies to Ecologies”
Atau Tanaka (Goldsmiths University of London)
- “Human-centered Multimodal Machine Intelligence”
Shrikanth (Shri) Narayanan (University of Southern California)

ピックアップ 論文紹介

視線

- 視線のふるまいからユーザの「困惑」を検出
A Neural Architecture for Detecting User Confusion in Eye-tracking Data
- Deepfake動画を見破るときの視線のふるまいの分析
The eyes know it: FakeET- An Eye-tracking Database to Understand Deepfake Perception

感情認認

- 生理指標（脳波や心拍数）を用いた感情認識
Is She Truly Enjoying the Conversation?: Analysis of Physiological Signals toward Adaptive Dialogue Systems
- 脳波と表情の信号からの感情推定
Multimodal Gated Information Fusion for Emotion Recognition from EEG Signals and Facial Behaviors (short)
- より自然な場面を想定したデータセット構築
 - (1) **MORSE: Multimodal sentiment analysis for Real-life Settings**
 - (2) **MSP-Face Corpus: A Natural Audiovisual Emotional Database**

マルチモーダル統合

- マルチモーダル統合におけるバイアス性を軽減する手法の検討
Mitigating Biases in Multimodal Personality Assessment
- 認知科学の知見に基づいたマルチモーダル統合モデル
Multimodal Data Fusion based on the Global Workspace Theory

コミュニケーションスキル

- 面接の形態（対面 vs. 非対面, 同期 vs. 非同期）の比較

Understanding Applicants' Reactions to Asynchronous Video Interviews Through Self-reports and Nonverbal Cues

- 「感情表現力」のマルチモーダルな推定

Toward Multimodal Modeling of Emotional Expressiveness

- プrezentationの質の評価と発表者へのフィードバック

Multimodal Assessment of Oral Presentations using HMMs (short)

- プrezentationの評価におけるプレゼン形態の違い（ビデオ, 音声, 書き起こし）の影響

Effect of Modality on Human and Machine Scoring of Presentation Videos (short)

- 面接における形態と文化の違いの影響を調査

Conventional and Non-conventional Job Interviewing Methods: A Comparative Study in Two Countries (short)

メンタルヘルス

- うつ病の度合いの推定

(1) Depression Severity Assessment for Adolescents at High Risk of Mental Disorders

(2) Detecting Depression in Less Than 10 Seconds: Impact of Speaking Time on Depression Detection Sensitivity

- 認知症の人の Engagement (タスクへの参与度合い) 認識

Towards Engagement Recognition of People with Dementia in Care Settings

独創的なタスク

- 「褒める」場面のマルチモダールデータコーパス
Analyzing Nonverbal Behaviors along with Praising (short)
- マルチモダールなユーモアの検出
Punchline Detection using Context-Aware Hierarchical Multimodal Fusion (short)
- 語学学習場面における学習者のUncertainty（未確信）の検出
Detection of Listener Uncertainty in Robot-Led Second Language Conversation Practice (short)
- 博物館のインタラクティブな展示において、ユーザの初期のふるまいからその展示の滞在時間を予測
Early Prediction of Visitor Engagement in Science Museums with Multimodal Learning Analytics
- 音声・視覚・言語の情報から表情（action unit）の強度を予測
Estimating the Intensity of Facial Expressions Accompanying Feedback Responses in Multiparty Video-Mediated Communication

コロナ禍特有？

- 自分の顔を触れたかどうかの推定
Analysis of Face-Touching Behavior in Large Scale Social Interaction Dataset
- マルチモーダルセンサを用いた咳イベントの検出
Automated Time Synchronization of Cough Events from Multimodal Sensors in Mobile Devices (short)

ICMI 2021 @Montreal, Canada

- 論文投稿締切り

2021年 5月 26日

- 会議開催期間

2021年 10月 18~22日



The image shows a panoramic view of the Montreal skyline at sunset, featuring the Ferris wheel and the St. Lawrence River.

23rd ACM International Conference on Multimodal Interaction
Montreal, Canada. October 18-22nd, 2021

The 23rd ACM International Conference on Multimodal Interaction (ICMI 2021) will be held in Montreal, Canada. ICMI is the premier international forum for multidisciplinary research on multimodal human-human and human-computer interaction, interfaces, and system development. The conference focuses on theoretical and empirical foundations, component technologies, and combined multimodal processing techniques that define the field of multimodal interaction analysis, interface design, and system development. ICMI 2021 will feature a single-track main conference which includes: keynote speakers, technical full and short papers (including oral and poster presentations), special sessions, demonstrations, exhibits and doctoral spotlight papers.

News and Updates

Paper Submission Deadline
May 26, 2021

Reviews Provided to Authors
July 7, 2021

Author Rebuttals Due
July 12, 2021

Notification of Acceptance
July 26, 2021

Camera Ready Papers Due
August 16, 2021

Conference Date
October 18-22, 2021

<https://icmi.acm.org/2021/index.php?id=home>



INTERSPEECH 2020
OCTOBER 25-29/ SHANGHAI, CHINA
SHANGHAI INTERNATIONAL CONVENTION CENTER

INTERSPEECH 2020
Cognitive Intelligence for Speech Processing



國際會議報告 INTERSPEECH 2020

京都大学 大学院情報学研究科
博士後期課程 1年
山本 賢太



INTERSPEECH2020 概要

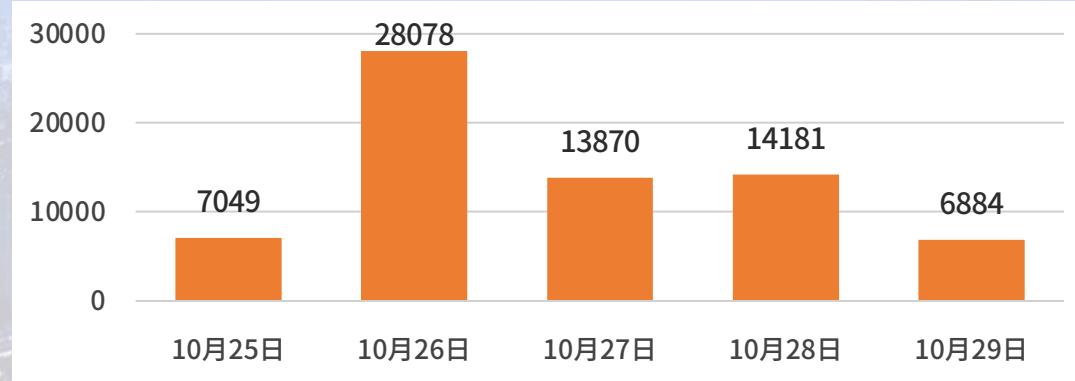
2

- ISCA (International Speech Communication Association) 主催
- 音声全般が対象
- 今年はオンライン開催（上海）
- 日程：2020年10月25日～10月29日（5日間）
初日はチュートリアルのみ
- 採択率：47 % (2019 年は 49%)
- 論文数：1021 件

オンライン開催

3

- 発表形式：全てビデオ発表に変更
 - 15分の発表ビデオと1分のハイライトビデオ
 - 当日は質疑応答のみ（Zoom）
- 参加人数：約 2000 人 (Registration)
- アクセス数：live broadcasts



Keynote

4

- “The cognitive status of simple and complex models”
Janet B. Pierrehumbert (University of Oxford)
- “Brain networks enabling speech perception in everyday settings”
Barbara Shinn-Cunningham (Carnegie Mellon University)
- “Doing Something we Never could with Spoken Language Technologies-from early days to the era of deep learning”
Lin-shan Lee (National Taiwan University)
- “Successes, Challenges and Opportunities for Speech Technology in Conversational Agents”
Shehzad Mevawalla (Amazon Alexa)

セッション

5

- 約 100 セッション
- 音声認識，音声合成がメイントピック
- 対話関連の研究に関連するセッションも複数ある

認識系が多め

対話研究に関係するセッション

Spoken Dialogue System

Conversational Systems

Spoken Language Understanding

Speech Emotion Recognition

ピックアップ論文

- “SpeechBERT: An Audio-and-text Jointly Learned Language Model for End-to-end Spoken Question Answering”
 - 単語ごとの音声をBERTの分散表現に近づくように学習
 - QAにおける文脈のテキストの代わりに音声をBERTベースのモデルへ入力
- “Dimensional Emotion Prediction based on Interactive Context in Conversation”
 - 話者と対話相手の両方の感情価と覚醒度を考慮した感情予測
 - 両話者の音響特徴量とテキスト特徴量を用いた階層構造のGRU
- “HRI-RNN: A User-Robot Dynamics-Oriented RNN for Engagement Decrease Detection”
 - リアルタイムでエンゲージメントの低下を予測
 - ユーザの画像・音声データに加えて、ロボットの音声データも利用

ピックアップ論文

7

- “Prototypical Q Networks for Automatic Conversational Diagnosis and Few-Shot New Disease Adaption”
 - 少数データによる対話マネージャーを学習（自動診断タスク）
 - 症例などのプロトタイプの分散表現を用いたマッチング
- “End-to-End Task-oriented Dialog System through Template Slot Value Generation”
 - テンプレート分類 + 知識検索 + スロット埋め込みをEnd-to-Endでモデル化
 - 発話のテンプレート生成を教師なし学習
- “Task-Oriented Dialog Generation with Enhanced Entity Representation”
 - 文脈情報と知識ベースの両方を用いたエンティティ表現の獲得
 - 単語のエンコードにおいて知識グラフ上で近い単語も利用

INTERSPEECH 2021

8

■ 開催地 : BRNO (CZECHIA, チェコ)

オンライン開催の可能性あり

■ 開催期間 : 2021年 8月 30日 ~ 2021年 9月 3日

■ 投稿締切 : 2021年 3月 26日



国際会議報告

EMNLP 2020

赤間 恵奈 (東北大)

人工知能学会 言語・音声理解と対話処理研究会 (JSAI SIG-SLUD 2020)
2020-12-01 @オンライン

基本情報

- Conference on Empirical Methods in Natural Language Processing (EMNLP)
 - 自然言語処理分野の権威ある国際会議のひとつ
 - Association for Computational Linguistics (ACL) が毎年開催
 - 採択論文は ACL Anthology (<https://www.aclweb.org/anthology/>) にて公開

Google Scholar Metrics (2020/11/30)

Categories > Engineering & Computer Science > Computational Linguistics ▾

Publication	<u>h5-index</u>	<u>h5-median</u>
1. Meeting of the Association for Computational Linguistics (ACL)	<u>135</u>	220
2. Conference on Empirical Methods in Natural Language Processing (EMNLP)	<u>112</u>	197
3. Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (HLT-NAACL)	<u>90</u>	148

EMNLP 2020 @Punta Cana, Dominican Republic Online



<https://2020.emnlp.org/>

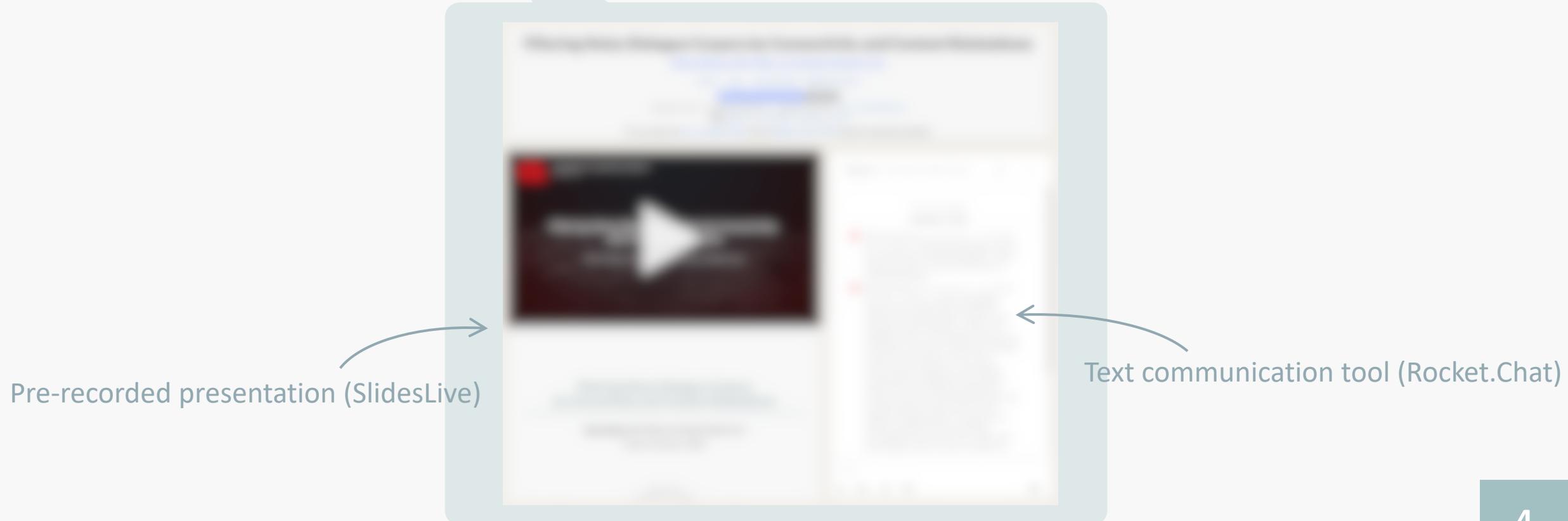
Schedule

- Anonymity period begins: May 1 (Preprint等の公開禁止)
- Submission deadline: June 3 (COVID-19と米デモで2回延びる; 5/11→6/1→6/3)
- **Author response period:** August 7-13
- Notification: September 14
- Main conference: November 16-18, Workshops (25) & Tutorials (7): November 19-20

約6ヶ月

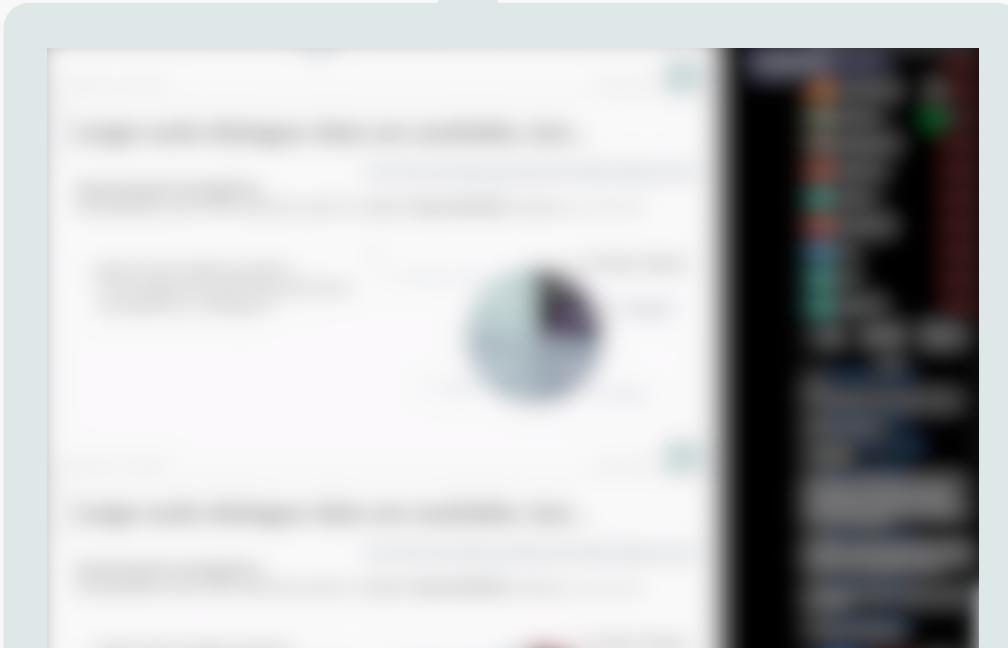
Main Conference の様子

- Presentation
 - Pre-recorded presentation (long 12 minutes, short 7 minutes)
 - Live Q&A session: Zoom (loral session; 1h) or Gather.town (poster session; 2h)

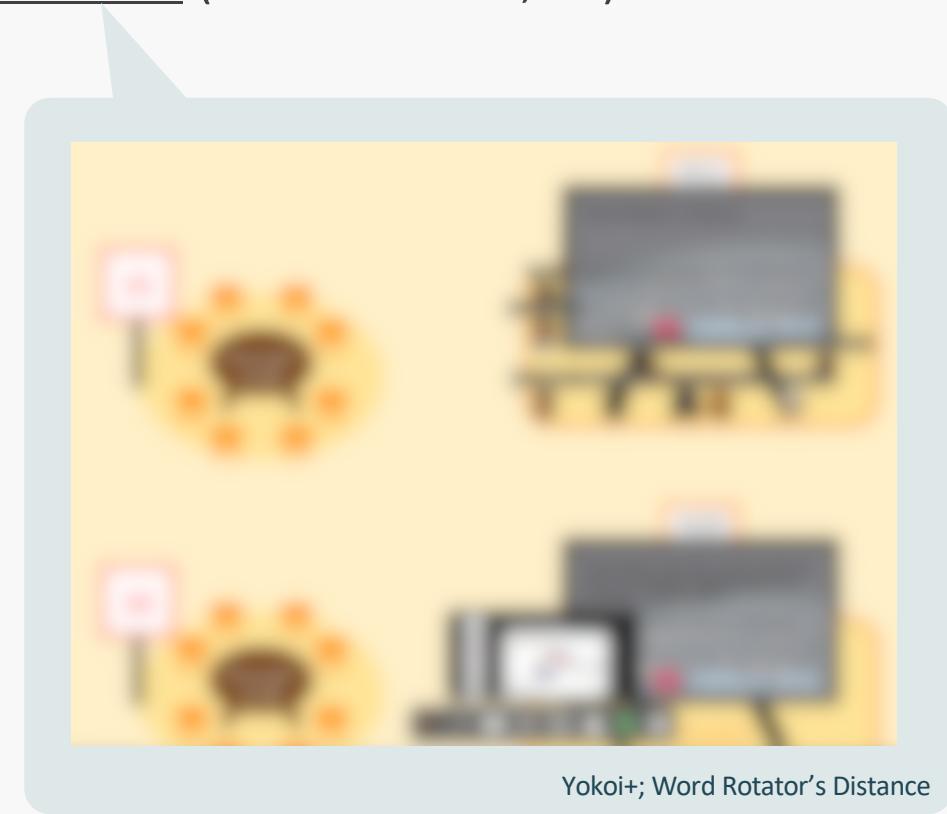


Main Conference の様子

- Presentation
 - Pre-recorded presentation (long 12 minutes, short 7 minutes)
 - Live Q&A session: Zoom (Oral session; 1h) or Gather.town (Poster session; 2h)



Akama+; Filtering Noisy Dialogue Corpora by Connectivity and Content Relatedness



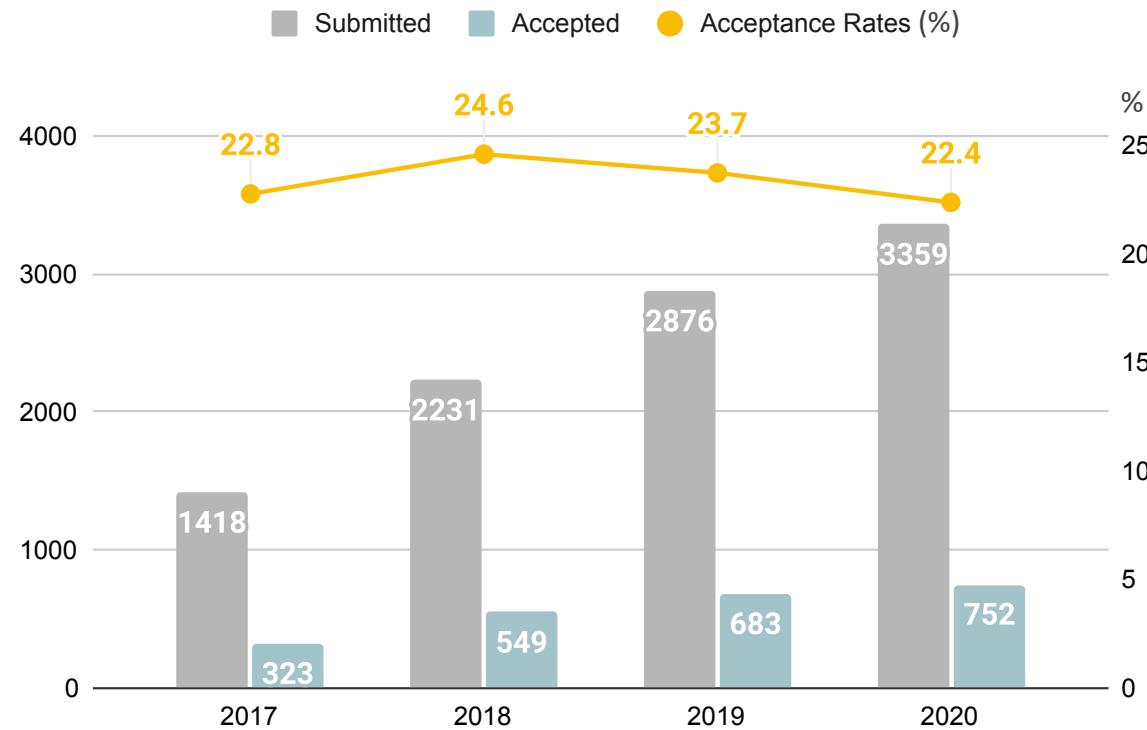
Yokoi+; Word Rotator's Distance

新しい試み

- Reproducibility Checklist の提出 提出時
 - Computing infrastructure, dataset (train/test split, download link), hyperparameter search, ...
 - これを元に査読者が reproducibility score をつける、採否の参考に
- 過去の査読結果・会議情報の添付 提出時
 - 査読結果を添付: 349 papers, 会議情報を添付: 855 papers
 - 本会議採択率 +6% (!)
- Findings
 - 惜しくも本会議に採択されなかった論文を別枠で採択 (520 papers, 投稿数の15.5%)
 - ACL Anthology で公開 a.k.a. "ACL archives"
- AAAI 2021 Fast Track
 - EMNLP 2020 でスコア 2.8 以上を獲得した long paper は、AAAI 2021 の (1st phrase を飛ばして) 直接 2nd phase の査読に進むことが可能

統計・傾向

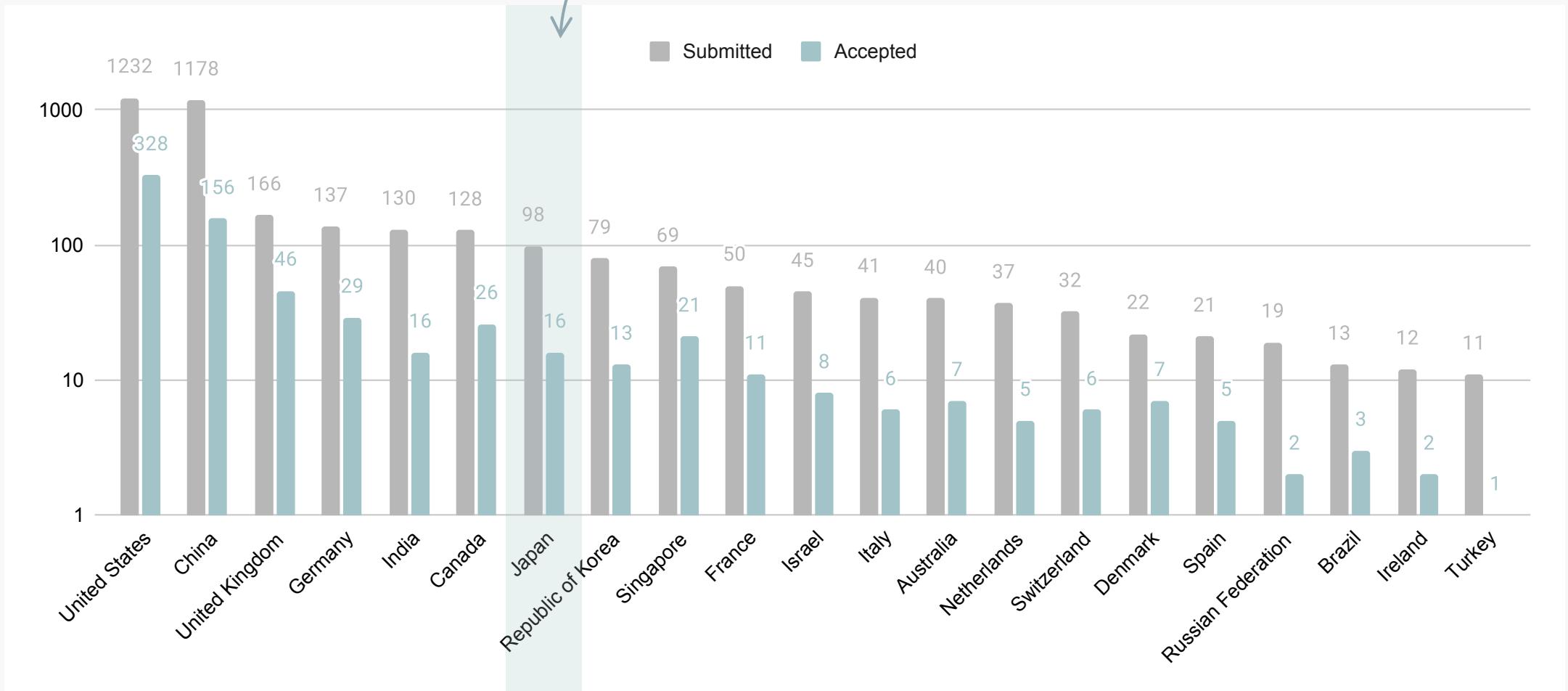
投稿数・採択数・採択率の推移



- EMNLP 2020 では
 - 投稿数 3,359, 採択数 752 (過去最大)
 - 採択率 22.4%
- **投稿数は増加傾向**
 - 2017→2020で2倍以上に
 - AI/NLP系トップ会議で同様の傾向
- **採択率は横ばい or 若干低下傾向**

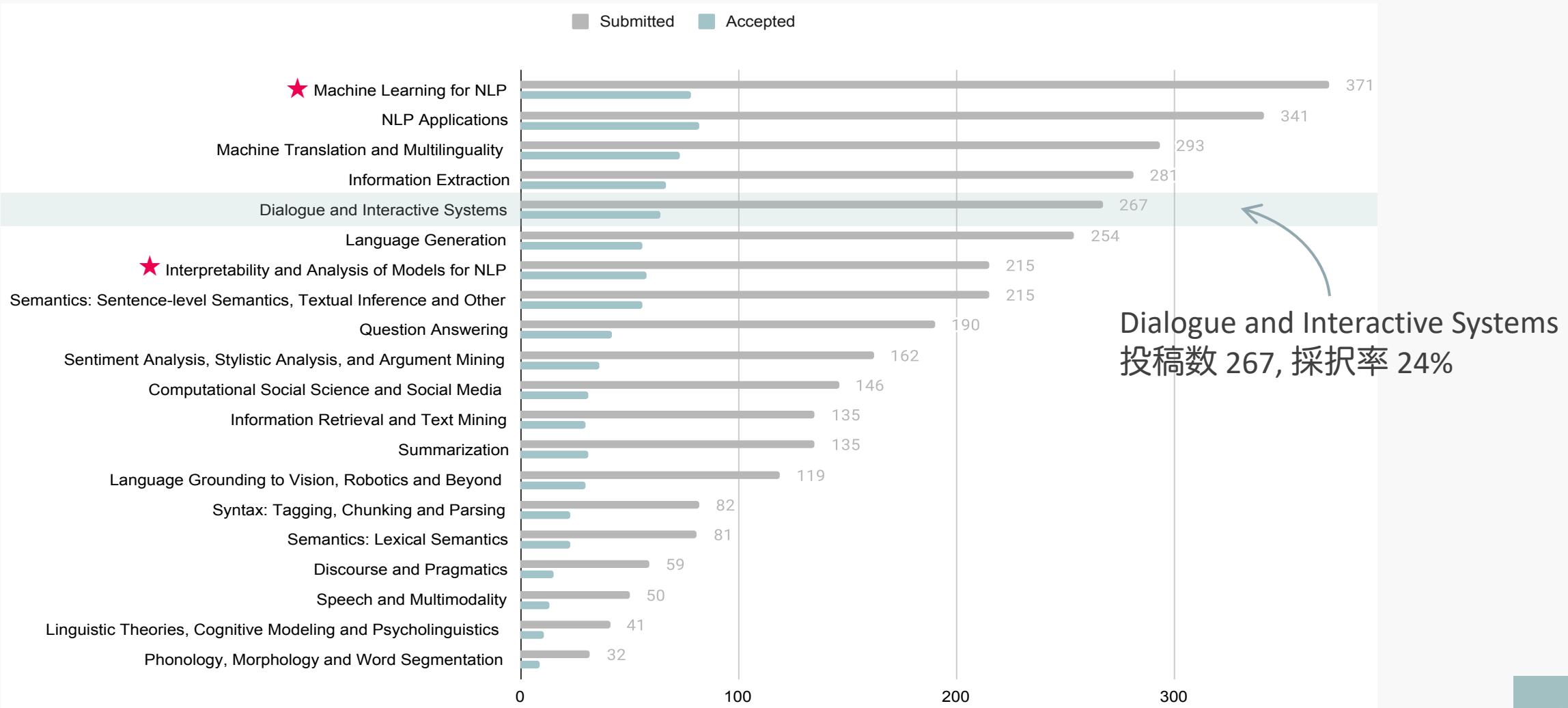
統計・傾向

国別投稿数・採択数



統計・傾向

エリア別投稿数 (全20エリア) エリア毎の採択率: 21~29%



対話関連研究

Dialogue and Interactive Systems area: **62 papers**

- **Neural dialogue system**
 - 学習戦略 [Ko+; Jaques+; Zhu+], モデルアーキテクチャ [Chen+; Ji+], 訓練データ [Zhang+; Akama+]
- **Task-oriented dialogue** (Language Generation area にもいくつか)
 - Andreas+; Task-Oriented Dialogue as Dataflow Synthesis
 - Lin+; MinTL: Minimalist Transfer Learning for Task-Oriented Dialogue Systems
 - Wu+; TOD-BERT: Pre-trained Natural Language Understanding for Task-Oriented Dialogue
- **New dataset**
 - Kamezawa+; A Visually-grounded First-person Dialogue Dataset with Verbal and Non-verbal Responses
 - Zeng+; MedDialog: Large-scale Medical Dialogue Datasets
 - Bastianelli+; SLURP: A Spoken Language Understanding Resource Package

論文紹介

- **Personalization vs Privacy**

- Kim+; Will I Sound Like Me? Improving Persona Consistency in Dialogues through Pragmatic Self-Consciousness
- Majumder+; Like hiking? You probably enjoy nature: Persona-grounded Dialog with Commonsense Expansions
- Song+; Profile Consistency Identification for Open-domain Dialogue Agents

論文紹介

- Personalization vs Privacy

Personal Information Leakage Detection in Conversations

Qiongkai Xu^{1,2}, Lizhen Qu^{3*}, Zeyu Gao¹, Gholamreza Haffari³

¹The Australian National University, Canberra, ACT, Australia

²Data61 CSIRO, Canberra, ACT, Australia

³Monash University, Clayton, VIC, Australia

- チャットボットを介した個人情報の漏洩は実社会では深刻な問題
- システムが生成した応答に個人情報が含まれているかを検出するタスクを提案
 - 評価セット PERSONA-LEAKAGE を作成（発話と個人情報の対応関係・情報の重要度）
 - 高度にペルソナ付与されたモデルは情報漏洩の潜在的なリスクを持つことを示唆

論文紹介

Best Paper Award (Honourable Mention Papers; 4/752)

Spot The Bot: A Robust and Efficient Framework for the Evaluation of Conversational Dialogue Systems

Jan Deriu¹, Don Tuggener¹, Pius von Däniken¹, Jon Ander Campos³,
Alvaro Rodrigo², Thiziri Belkacem⁴, Aitor Soroa³, Eneko Agirre³, and Mark Cieliebak¹

再現性がある 所要時間&ワーカ負担減

- 頑健で時間効率の良いシステム評価の方法論
 - 良いシステムは人間をよく模倣する
 - **Bot A-bot B 会話を人手アノテーション 「A/B は {human, unsure, bot} ?」**
→ Win-rate (*human > unsure > bot*) によりシステムを順位付け
- 面白い知見
 - **Sensibleness** が評価に大きく影響 (vs Fluency, Specificity [Adiwardana+'20])
 - Blender [Roller+'20] 高評価

EMNLP 2021 @Punta Cana, Dominican Republic (hopefully...)



EMNLP 2021

The 2021 Conference on Empirical Methods
in Natural Language Processing

7th – 11th November 2021

<https://2021.emnlp.org/>

Schedule

- Submission deadline: **late May or early June, 2021**
- Conference: **November 7-11, 2021**