

相乗り環境における会話支援AIが
利用者の会話行動・会話内容・利用者体験に
与える影響の実証研究

24W6064A 藤森美帆

NAISTモビリティオーケション



- ・乗り捨て可能なシェアリング
- ・駐車場
NAISTキャンパス内
株式会社国際電気通信基礎技術
研究所(ATR)
- ・対象ユーザ
NAISTの学生・教職員
ATRの社員

NAISTモビリティオーケションの課題

課題1: オークションで勝利したユーザしか車両を利用できないため車両効率が悪い

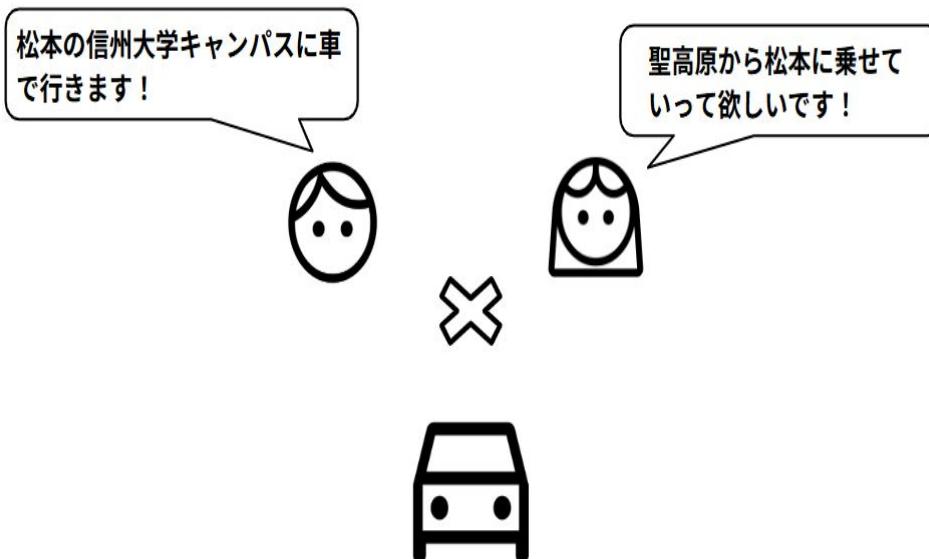
課題2: 運転免許を持たないユーザはモビリティオーケションに参加できない

以上の課題を解決するため、

相乗りの仕組みを導入することを考えている

相乗りとは

出発地・目的地や出発時刻が近い利用者同士をマッチングし、同一車両で移動を共有する交通サービス



- ・車両1台に対し複数人ユーザが乗って移動できる

- ・運転免許を持たないユーザでも移動できる

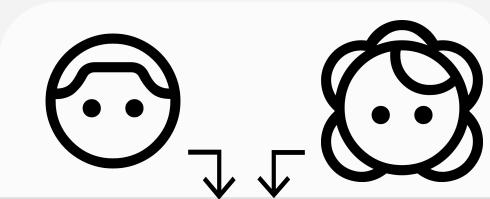


NAISTモビリティオークションにおける課題解決が可能

相乗りにおける課題の1つ

相乗り車内＝閉塞感を感じやすい空間

- ユーザー同士の距離が近い
- 途中で退出できない



同乗相手が初対面同士だと…

- 沈黙の継続が気まずい
- 会話の起こりづらさを感じる

例: Sarriera ら [1] における動的ライドシェア(UberPool／Lyft Line 等)を対象とした利用者調査では、同乗者と同乗した際に「気まずい沈黙(awkward silences)」が少なくとも半分の頻度で生じたと回答したユーザが48%存在する。

相乗り車内の「気まずさ」の解決策

【通常の会話】

先行研究: AIが質問・話題提示で初対面者の会話を支援

例: Nguyenら[2]は、事前に得た参加者情報に基づいてAIが話題候補を提示することで、初対面者間の会話を支援する手法を提案している

【車内空間の会話】

車内空間のような、近接性と退出困難性が強い環境におけるAIの会話支援は困難とされ、介入の設計指針や評価知見は十分に整理されていない。



車内における会話のおこりづらさ、沈黙の継続を AIの会話支援で改善する

[2]Nguyen, D. T., Iqbal, S. T., & Ofek, E. (2015), "The Known Stranger: Supporting Conversations between Strangers with Personalized Topic Suggestions"

本研究の目的とアプローチ

相乗りにおける会話のおこりづらさ、沈黙の継続を改善するために、
利用者の会話行動・利用者体験・会話内容から会話支援 AIの有効性を検証する

目的達成のため、本研究では以下のリサーチクエスチョンを設定し、会話支援 AIを用いた実証実験を行った

RQ1: 会話支援AIは利用者の会話行動にどのような影響を与えるか

RQ2: 会話支援AIは利用者体験にどのような影響を与えるか

RQ3: 会話支援AIは利用者の会話内容にどのような影響を与えるか

会話支援AIエージェントの構築

会話の起こりづらさや沈黙時間の減少を目指す会話支援 AIエージェント



AIとリアルタイムで会話
スマホから利用できる
Webアプリケーション

目的地や周辺情報に基づい
た話題提供を行う
(GPS)

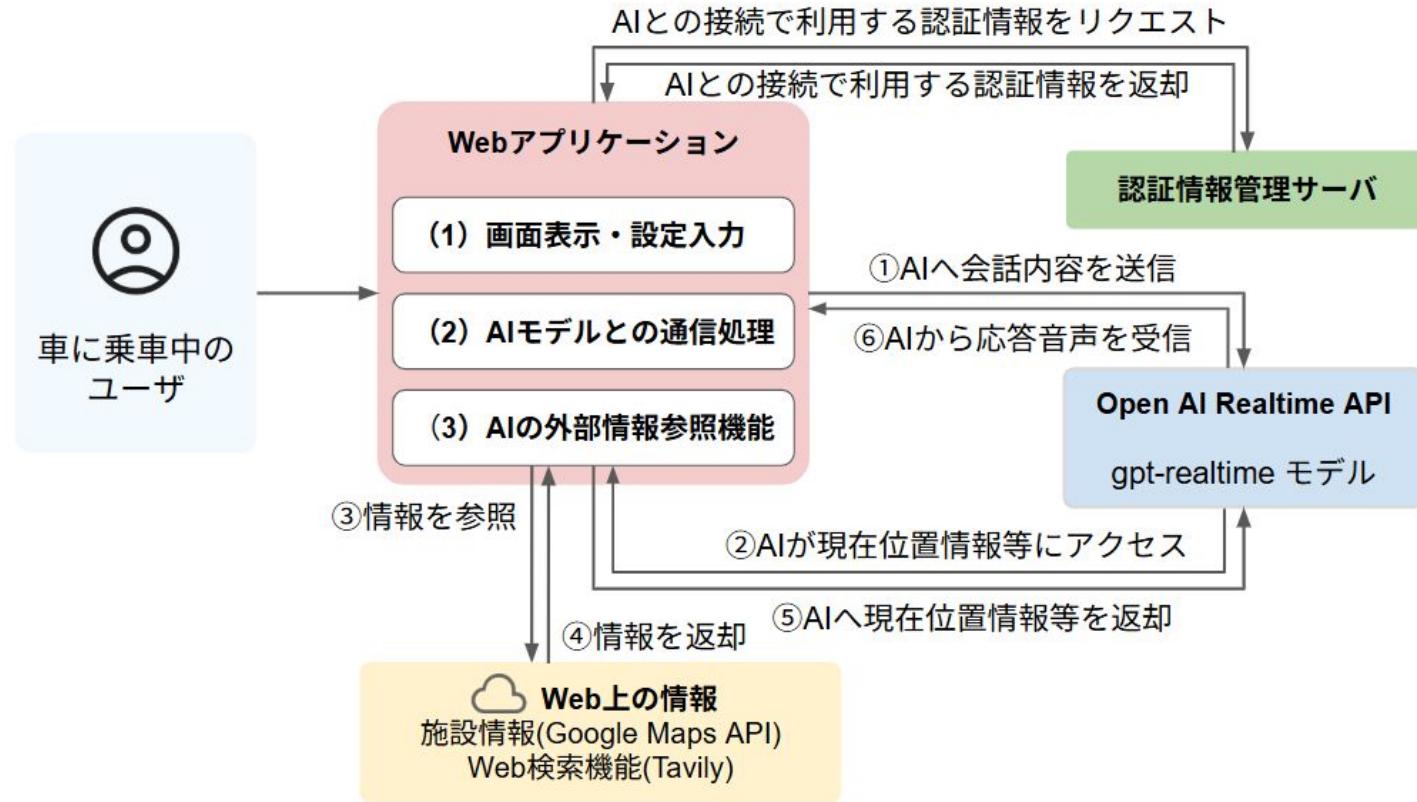
通常対話モード
初対面のユーザに共通の話
題を提供

ユーザ情報に基づいた話題
提供
(趣味、過去の経験、行ってみた
い国等)

会話中は割り込まない
沈黙が6秒続いたら
AI側から話題提供

ゲームモード
「3つのうち1つだけウソゲー
ム」を実施

会話支援AIエージェントの構成



評価実験

事前アンケート

- 参加者情報
- ビックファイブ
性格診断
- ゲームで扱う
項目

実証実験

- ・計9ペア(16人)
(2人の参加者が異なるペアで2回参加)
- ・停車中の車内
- ・運転手役1名前席、参加者2名は後部座席
- ・Plaud Note Pro(音声録音デバイス)で会話音
声を録音
- ・AIエージェントはスマホ画面で提示
- ・AIなし自由対話セッション(5分)
- ・AIあり自由対話セッション(5分)
- ・AIなしゲームセッション(5分)
- ・AIありゲームセッション(5分)

事後アンケート

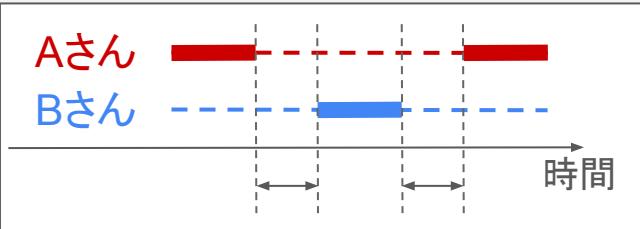
- 会話の質
- 親近感
- 相手理解
- プライバシー
懸念
- 話題
- 有用性
- 好感度

自由会話の想定

トピックを持った会話

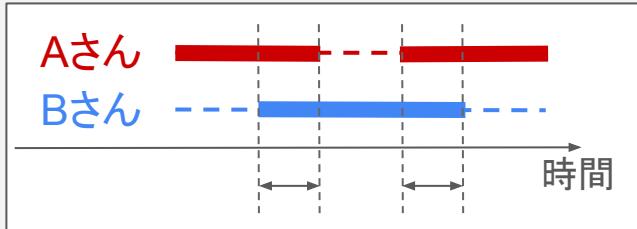
収集データ

沈黙時間・沈黙回数



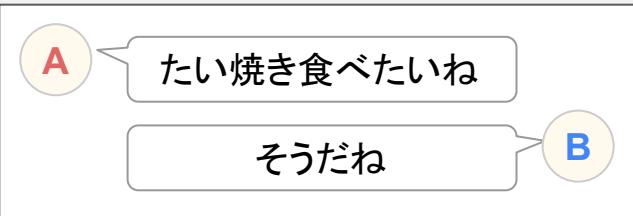
1秒以上の完全無音区間を「沈黙」として集計

同時発話回数



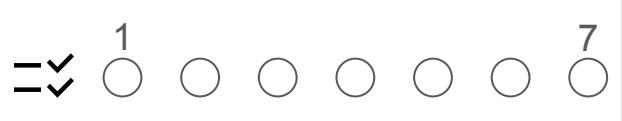
0.3秒以上の重複区間を「同時発話」として集計

追従発話回数



直前の発話に同意する応答を
「追従発話」として集計

性格特性・事後アンケート



事後アンケートから利用者体験のスコアと
参加者の性格特性「外交的」「内向的」「中間」を
得た

RQ1:会話支援AIは利用者の会話行動にどのような影響を与えるか

会話支援AIが参加者の会話行動にどのような影響を与えたかの分析を行う。
分析には実証実験で得られた以下のデータを用いる。以下のデータは会話行動を定量的に測る指標として先行研究で示されている。[3,4,5]

- ・**沈黙回数** : 1秒以上の完全無音区間
- ・**沈黙時間** : 完全無音空間の時間
- ・**同時発話回数** : 複数発話者における0.3秒以上の重複区間
- ・**追従発話回数** : 直前の発話に同意する応答(相づち、復唱、補完)

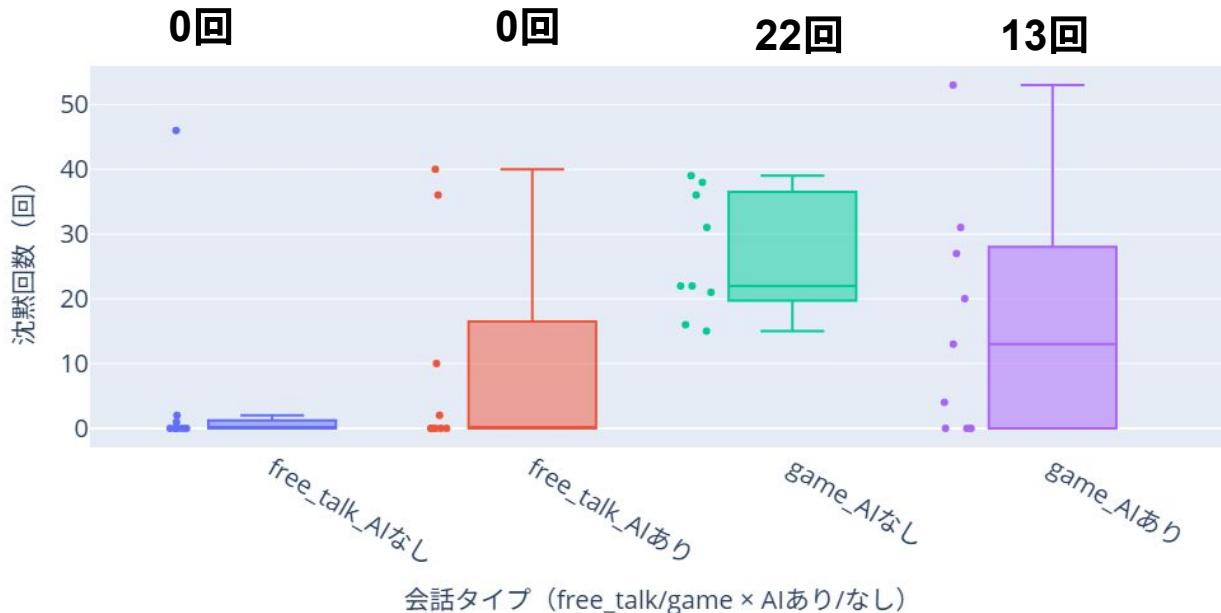
[3]Chowdhury, S. A., Stepanov, E., Danieli, M., & Riccardi, G. (2017) "Functions of Silences towards Information Flow in Spoken Conversation"

[4]Sacks, H., Schegloff, E. A., & Jefferson, G. (1974) "A Simplest Systematics for the Organization of Turn-Taking for Conversation"

[5]豊田 薫, 宮越 喜浩, 山西 良典, & 加藤 昇平(2012)"発話状態時間長に着目した対話雰囲気推定 "

RQ1の結果(沈黙回数、短い沈黙)

沈黙秒数が1秒以上4秒未満の長さにおける沈黙回数 (free_talk/game × AIあり/なし)



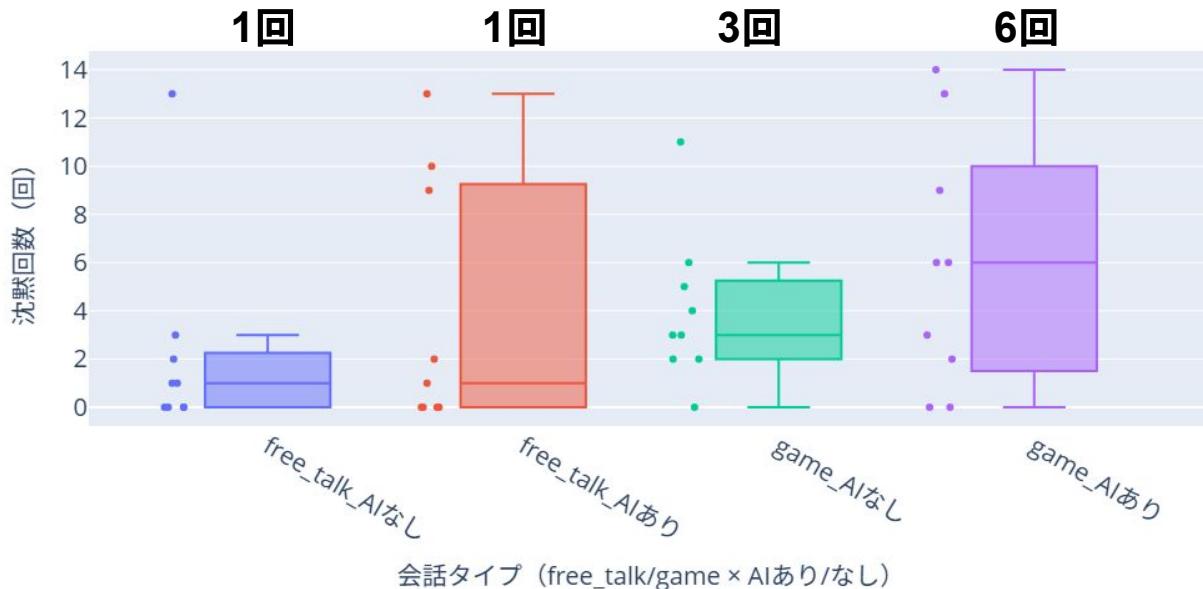
各会話タイプでみられた1秒以上4秒未満の沈黙回数

free talkセッションでは AI ありとなしで沈黙回数の中央値は変化しない

gameセッションでは AI ありで沈黙回数の中央値が減少する

RQ1の結果(沈黙回数、長い沈黙)

沈黙秒数が4秒以上10秒未満の長さにおける沈黙回数 (free_talk/game × AIあり/なし)



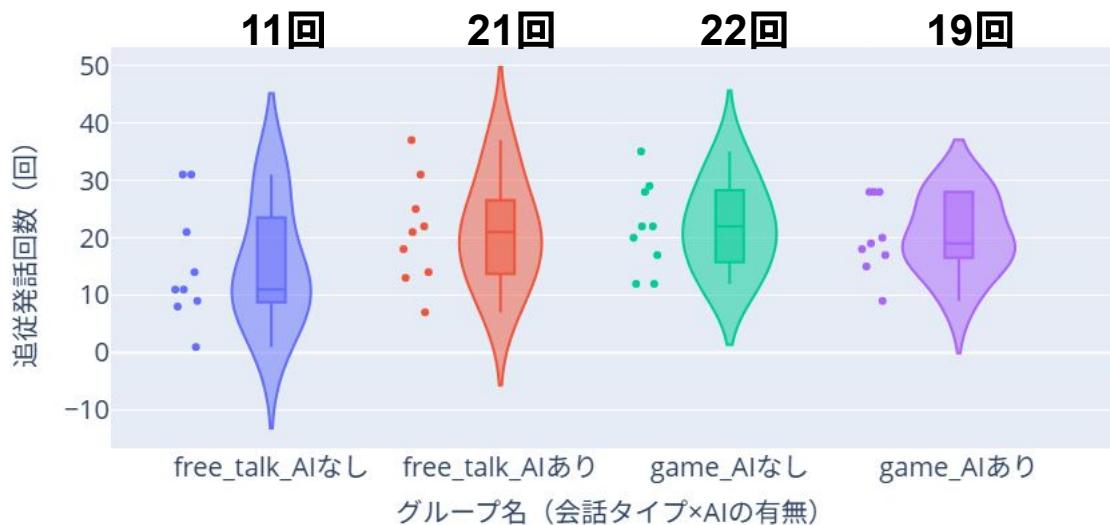
free talk / game とともに
AIありで沈黙回数が高 い
値をとるセッション が観察

gameセッションでは AI あ
りで沈黙回数の中央 値が
増加する

各会話タイプでみられた4秒以上10秒未満の沈黙回数

RQ1の結果(追従発話回数)

追従発話回数 (会話タイプ×AIのあり/なし)



free talkセッションでの中央値はgameセッションと比べ高くなっている

gameセッションでの中央値はfree talkセッションと比べ下がっている

各会話グループで見られた追従発話回数

RQ1考察

沈黙回数・沈黙時間

②game: 短い沈黙が減少→会話のテンポが速まった

③game: 長い沈黙が増加→AIがgameを補助することで参加者の「考える時間」が増えた

同時発話回数

free talk / game : AIなしの方がAIありより同時発話回数が多い

④AI介入が発話タイミングに影響し、同時発話を抑制した

⑤話者交代が安定し、参加者間の発話の偏りが小さくなつた

追従発話回数

free talk: AIありの方がAIなしより追従発話回数が多い

⑥参加しやすさが高まり、発話の偏りが小さくなつた

game: AIなしの方がAIありより追従発話回数が少ない

⑦AIが進行を担うことで、参加者同士のやり取りが減少した

→結果として「会話に参加しにくい」状態が生じた

RQ2:会話支援AIは利用者体験にどのような影響を与えるか

実証実験で行なった実験後アンケートから**利用者体験のスコア**を測定した。先行研究であるNguyenら[2]の事後アンケートThe Known Strangerで用いられた12項目)の質問紙を採用した

各アンケート項目における数値を得点とする(1点から7点の7点法)

会話の質(Q1–Q5)、親近感(Q6–Q8):各項目の平均値を得点にする

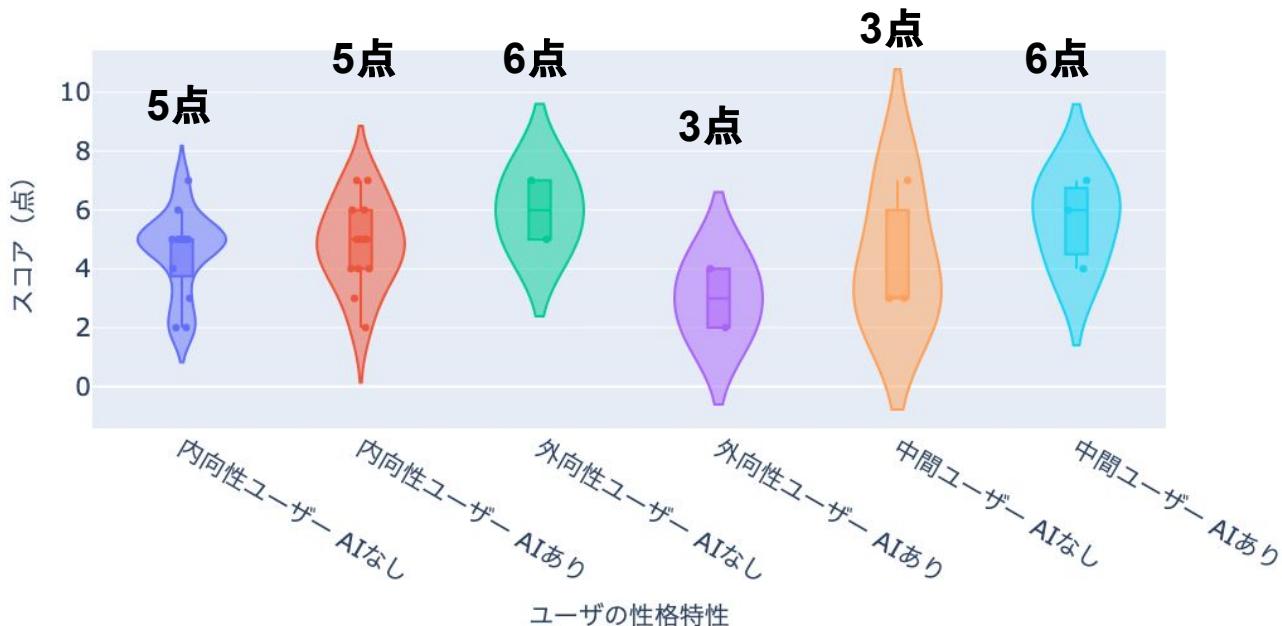
プライバシー懸念／相手理解／提示トピック有用性

セッション好感度(Q9–Q12):単項目の数値を得点にする

[2]Nguyen, D. T., Iqbal, S. T., & Ofek, E. (2015), "The Known Stranger: Supporting Conversations between Strangers with Personalized Topic Suggestions"

性格特性別における提示トピック有用性

事後アンケート結果：提示されたトピック提案の有用性



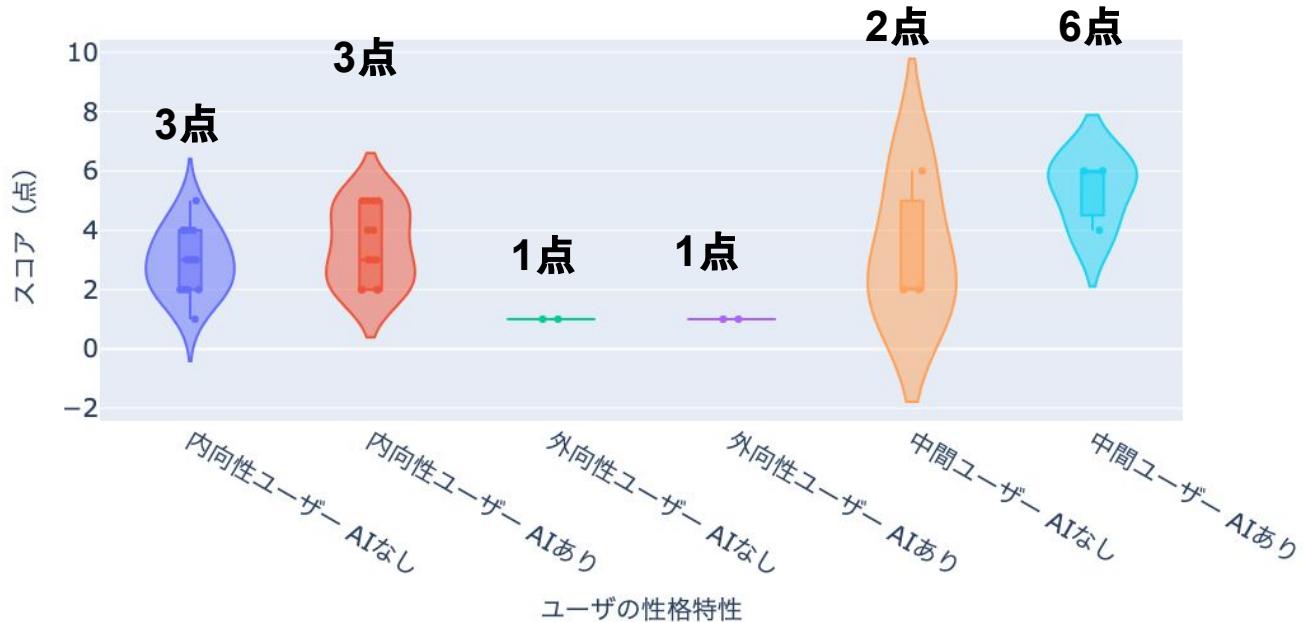
各性格特性でみられた提示トピック有用性におけるスコア

外向性ユーザにおける
中央値はAIなしの方が
AIありに比べて3点高い

中間ユーザにおける中
央値はAIありの方がAI
なしに比べて3点高い

性格特性別におけるプライバシー懸念

事後アンケート結果：プライバシーへの懸念



各性格特性でみられたプライバシーへの懸念におけるスコア

中間ユーザにおけるプライバシー懸念の中央値がAIありがAIなしより4点高い

外向性ユーザにおけるプライバシー懸念はAIありとなしで1点である

RQ2考察

トピック提案における有用性

- ⌚→外向性ユーザにおいては提案が邪魔になってしまったと思われる
- ⌚→中間ユーザにおいては会話の話題を探す負担を減らせた

プライバシーへの懸念

- ⌚→AIの介入で車内が監視されているような雰囲気ができ、閉鎖感や緊張感が高まってしまった

- ⌚→AIの介入が自然な会話の邪魔になってしまった
- ⌚→AIの介入で参加者同士の会話時間が短くなってしまった

RQ3:会話支援AIは利用者の会話内容にどのような影響を与えるか

実証実験で得られた録音記録の書き起こしを行い、テキストから話題の推移を可視化する。話題の推移から会話支援AIが参加者の会話内容にどのような影響を与えるかを見る。

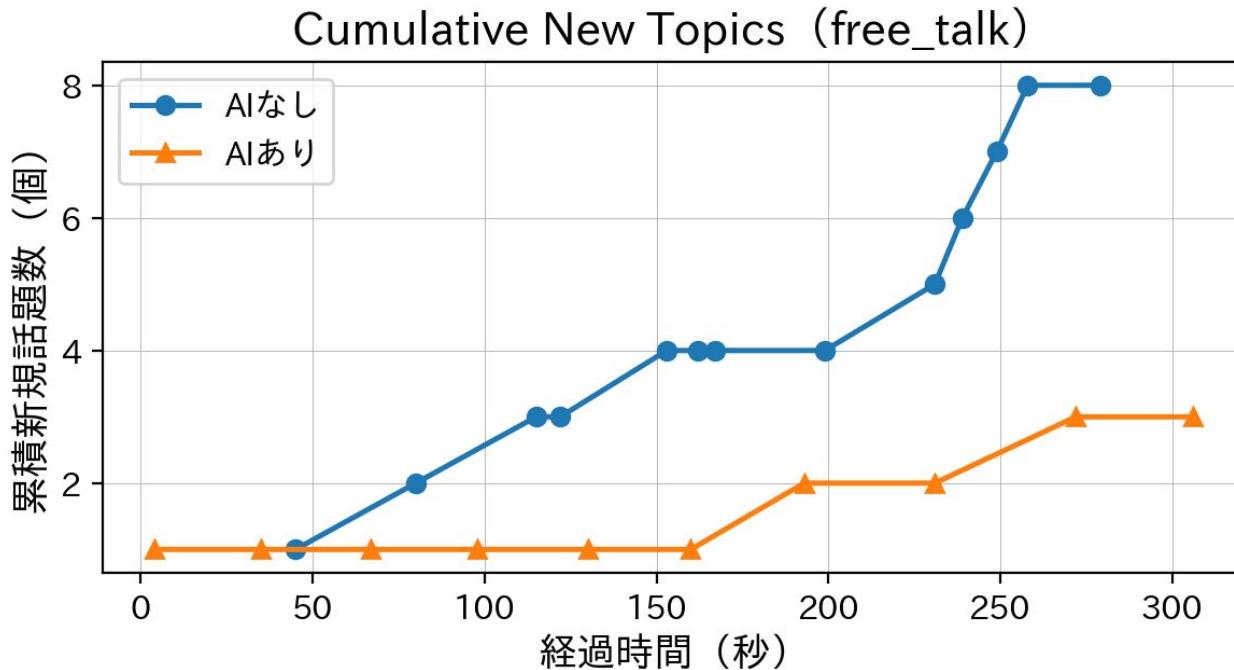
テキストの可視化では以下のグラフを用いた

Cumulative Topic-Depth(累積話題深度): 話題の深掘り度(階層の深さ)の推移を示す

Topic Lanes(話題レーン): 時系列上で話題ID(レーン)がどのように推移したかを示す

Cumulative New Topics(累積新規話題数): 新規話題が導入された回数を累積で示す

新規話題を抑制し話を深掘りしたケース



ペアAでみられた新規話題数の推移

AIありにおける新規話題数はAIなしの新規話題数と比較して値が一定

RQ3:考察

free talk:話の深掘り

ⒶAIありで新規話題が抑制される→同じ話題について深い話題になりやすい

game:進行の安定化・話題遷移の整理

ⒷAIありでTopic-DepthやLanesが比較的安定する例→司会進行・質問支援で話題遷移が整理される

まとめ

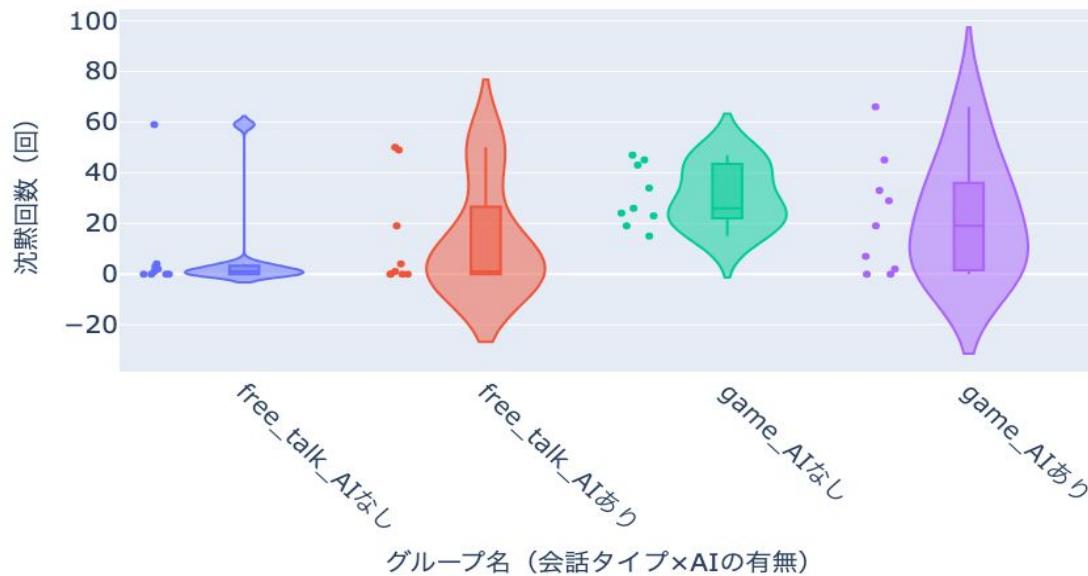
- ・話題作りができた
- 😊 ・会話参加の偏りを改善
- ・会話内容の深掘りができた
- ・会話テンポの向上

- ⌚ ・プライバシーへの懸念がある
- ・会話の質や相手理解における利用者体験が下がる

AIの介入が参加者の会話行動を変化させたことが分かった。一方でAIの介入で車内の閉塞感や会話の楽しさ、会話の質ではAIが邪魔になったことが分かる。

総沈黙回数

沈黙回数（回）（会話タイプ×AIのあり/なし）



free talk

AIありの中央値: 1回

AIなし: 1回

game

中央値は AIなし: 24回

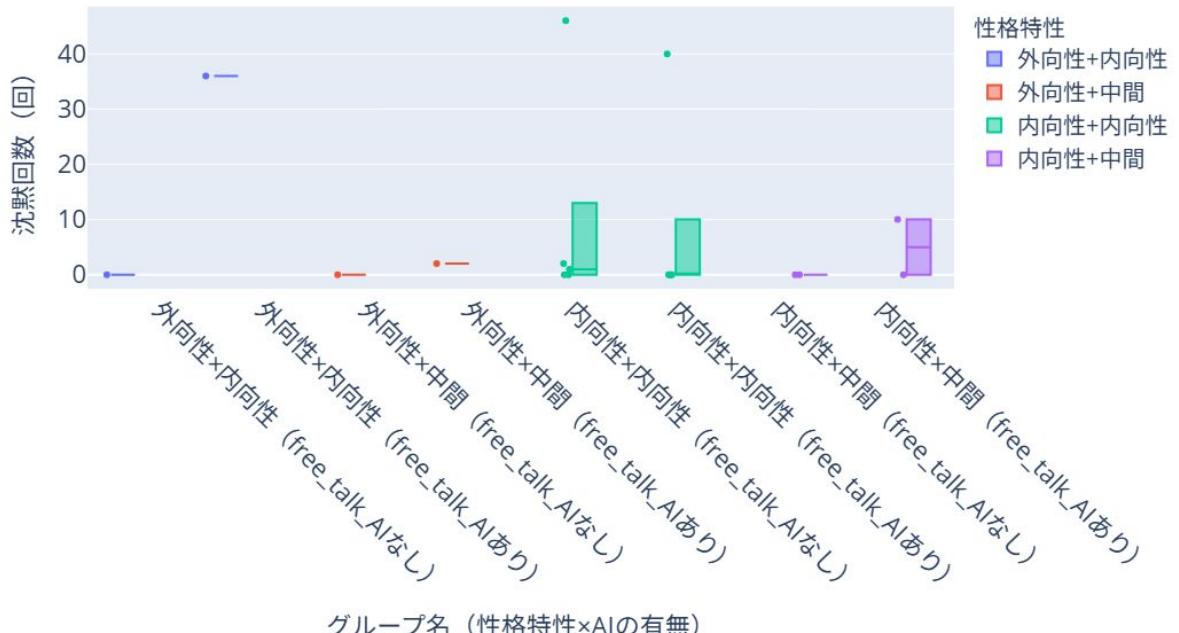
AIあり: 18回

AIありの方が沈黙回数が少ない

総沈黙回数

free talk時における沈黙回数、性格特性別の短い沈黙

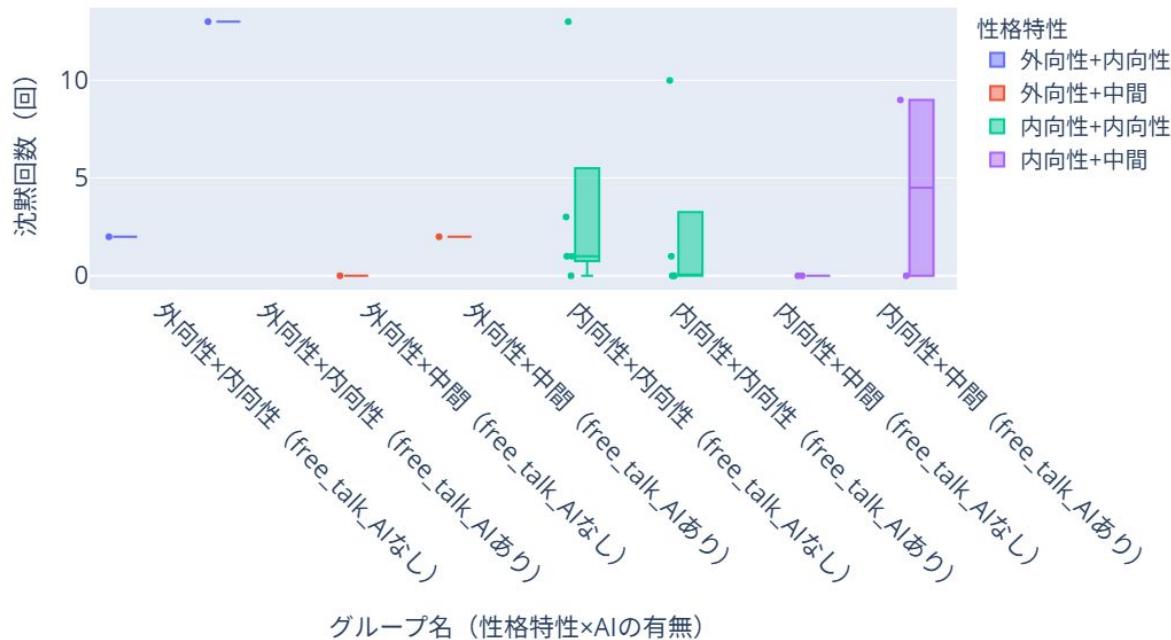
free_talk時において沈黙秒数が1秒以上4秒未満の長さにおける沈黙回数
(性格特性×AIあり/なし)



- ・外向性×内向性のペアにおいて沈黙回数の中央値がAIありで大幅に増加
- ・外向性×内向性、内向性×内向性でのペアではAIありとなしで中央値がほぼ横ばい

free talk時における沈黙回数、性格特性別の長い沈黙

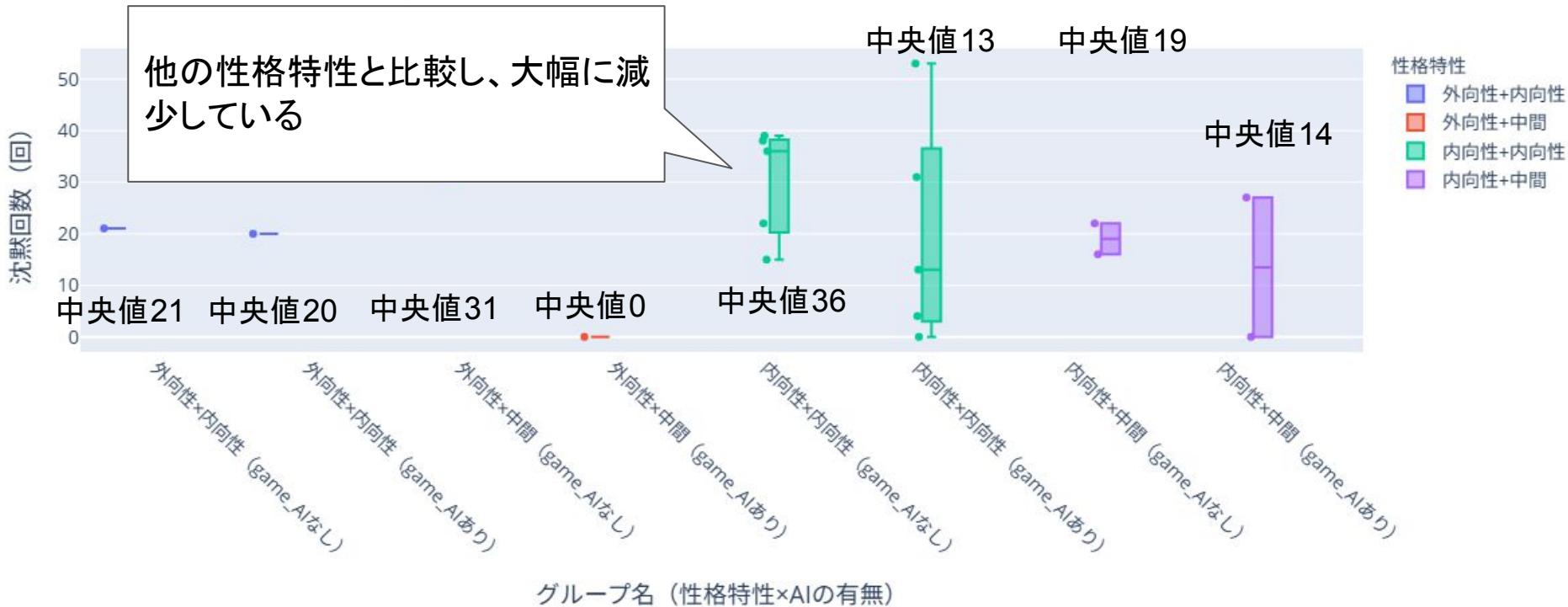
free_talk時において沈黙秒数が4秒以上10秒未満の長さにおける沈黙回数
(性格特性×AIあり/なし)



- ・外向性×内向性、内向性×中間において長い沈黙の中央値が大幅に増加
- ・内向性×内向性のペアにおける中央値はほぼ一定

game時における沈黙回数、性格特性別の短い沈黙

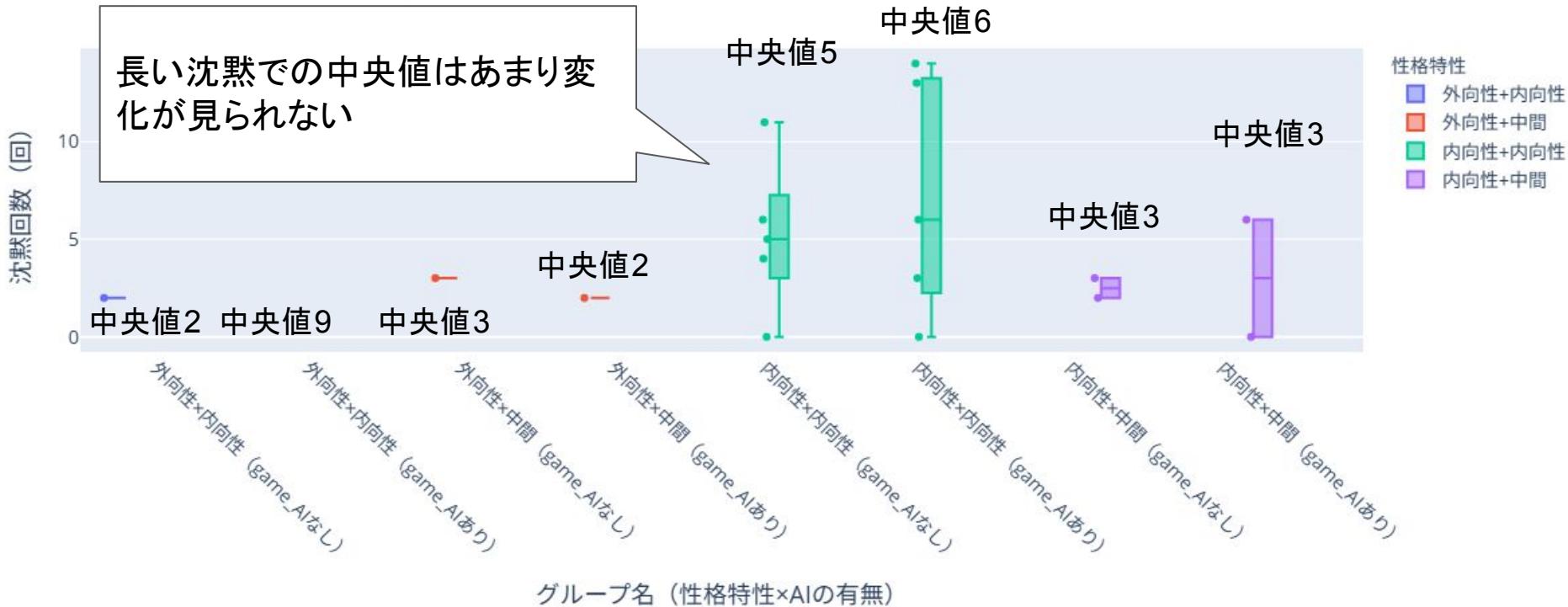
game時において沈黙秒数が1秒以上4秒未満の長さにおける沈黙回数（性格特性×AIあり/なし）



性格特性別でみられた1秒以上4秒未満の沈黙回数

game時における沈黙回数、性格特性別の長い沈黙

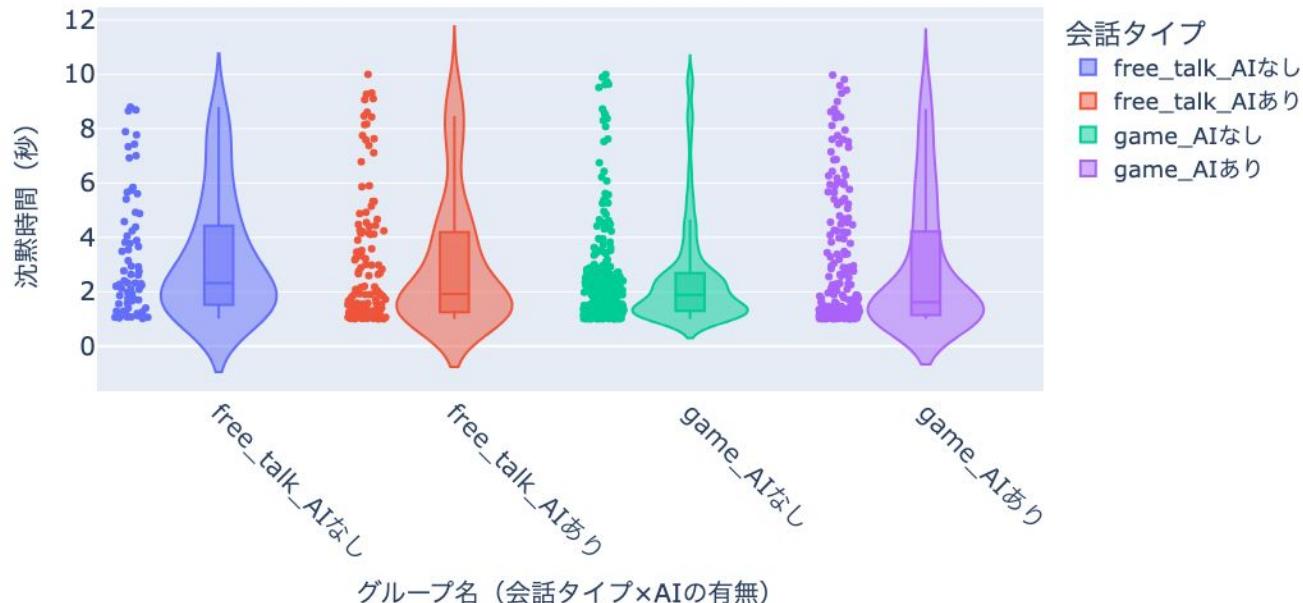
game時において沈黙秒数が4秒以上10秒未満の長さにおける沈黙回数（性格特性×AIあり/なし）



性格特性別でみられた4秒以上10秒未満の沈黙回数

総沈黙時間

沈黙時間 (秒) (会話タイプ×AIのあり/なし)

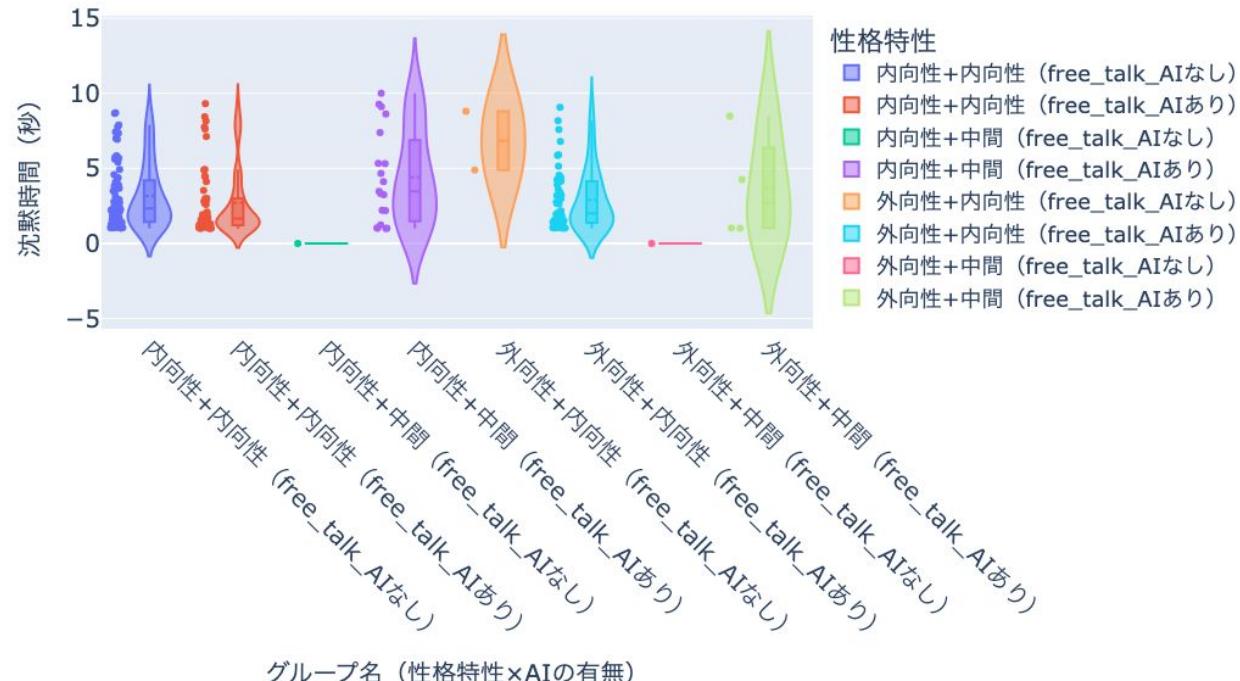


4秒を境にしてAIありでAIなしと比較して短い沈黙が減り、長い沈黙が増加している

各会話タイプでみられた沈黙時間

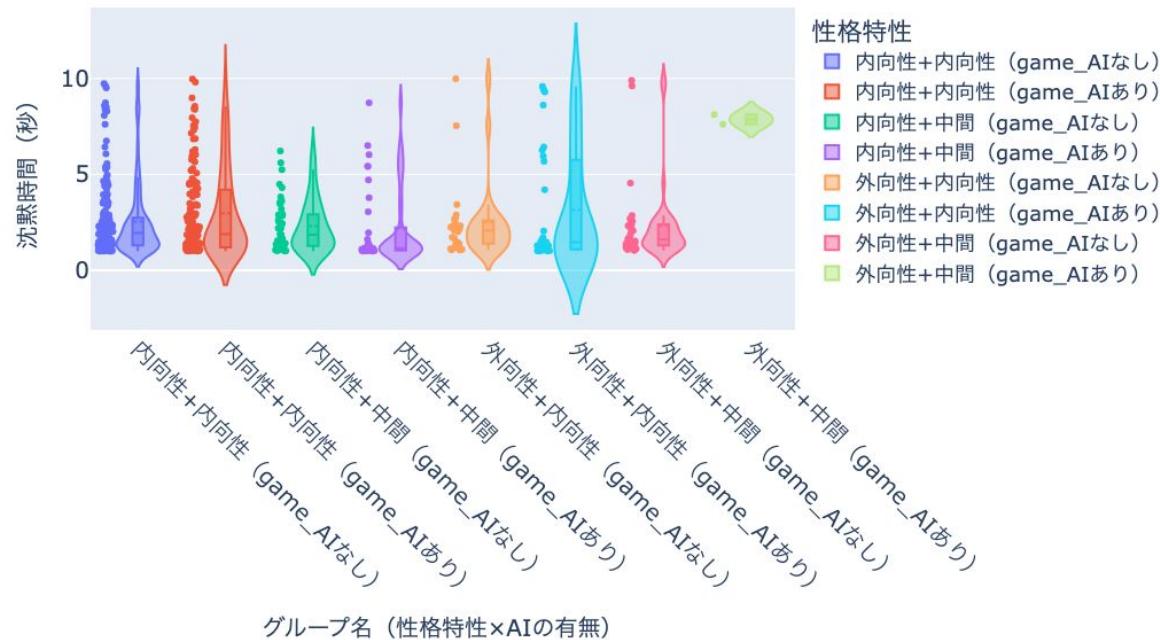
free talk時の性格特性別における沈黙時間

free_talk時の沈黙時間（秒）（性格特性×AIのあり/なし）



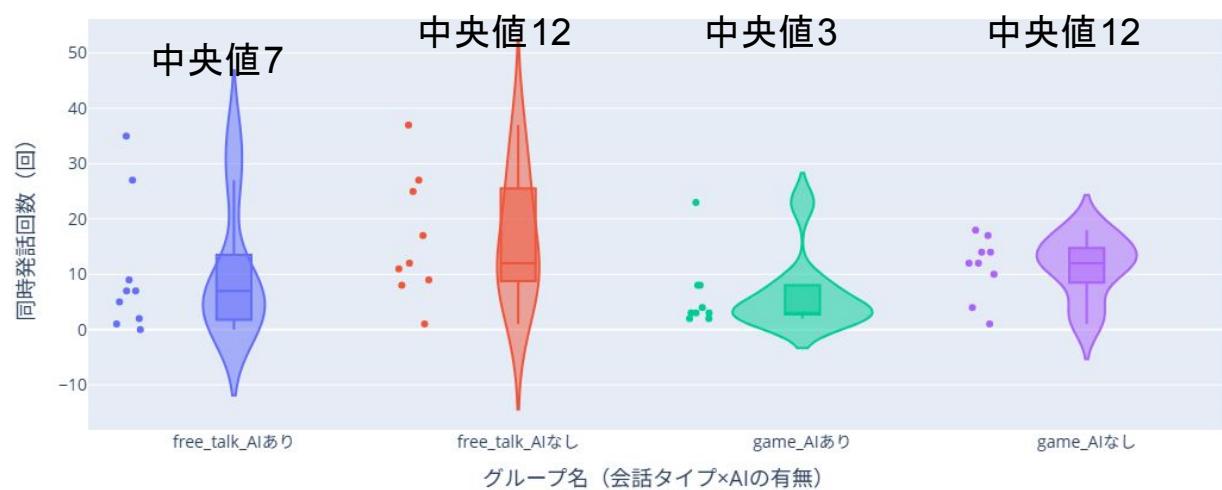
game時の性格特性別における沈黙時間

game時の沈黙時間（秒）（性格特性×AIのあり/なし）



総同時発話回数

同時発話回数（会話タイプ×AIのあり/なし）



free talk

AIなし: 同時発話回数が高い傾向

AIあり: 中央値が低下(低い方向に集中)

game

AIなし: 同時発話回数が高い

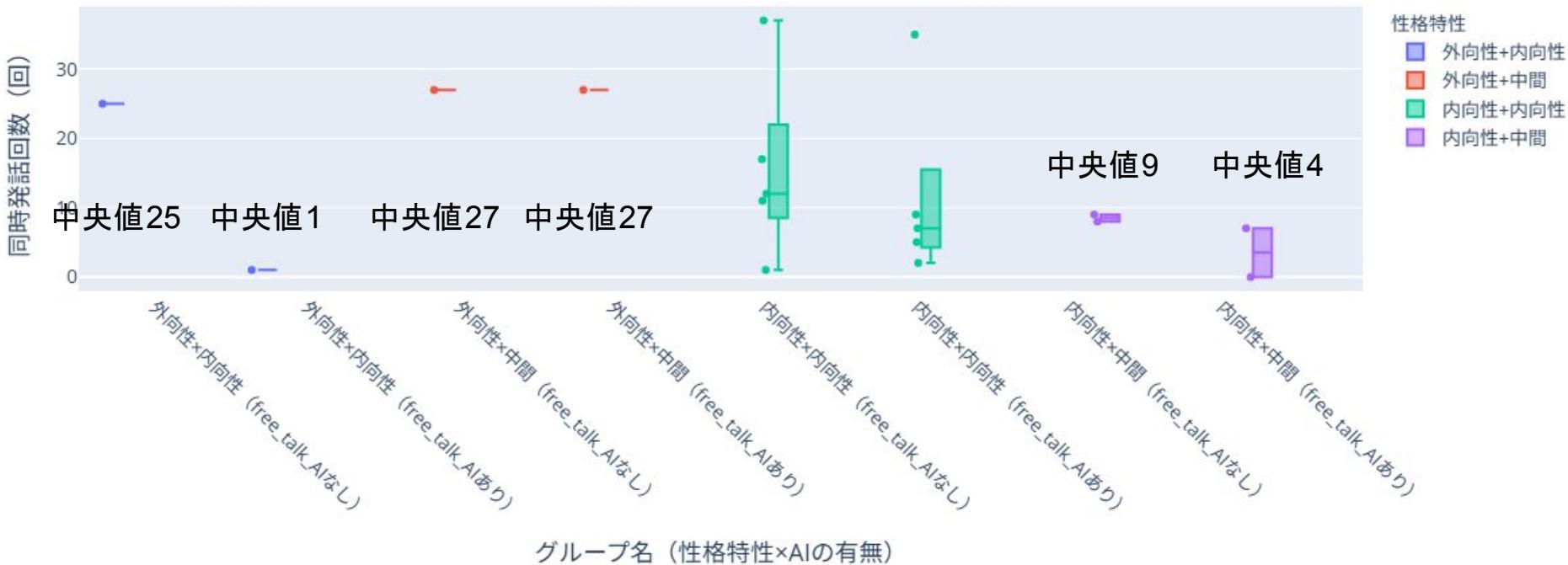
AIあり: 低い値に集中する傾向

各会話グループで見られた同時発話回数

free talk時における性格特性別の同時発話回数

free_talk時における同時発話回数(性格特性×AIあり/なし)

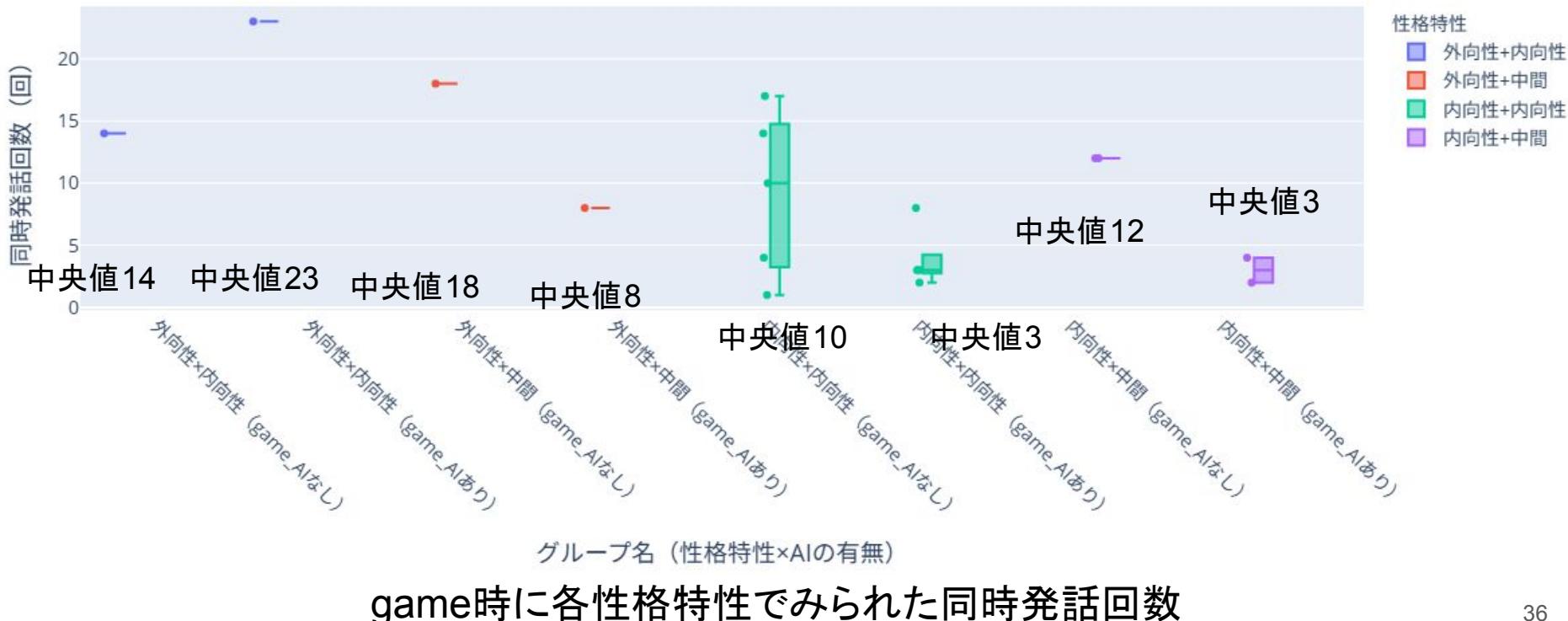
中央値12 中央値7



free talk時に各性格特性でみられた同時発話回数

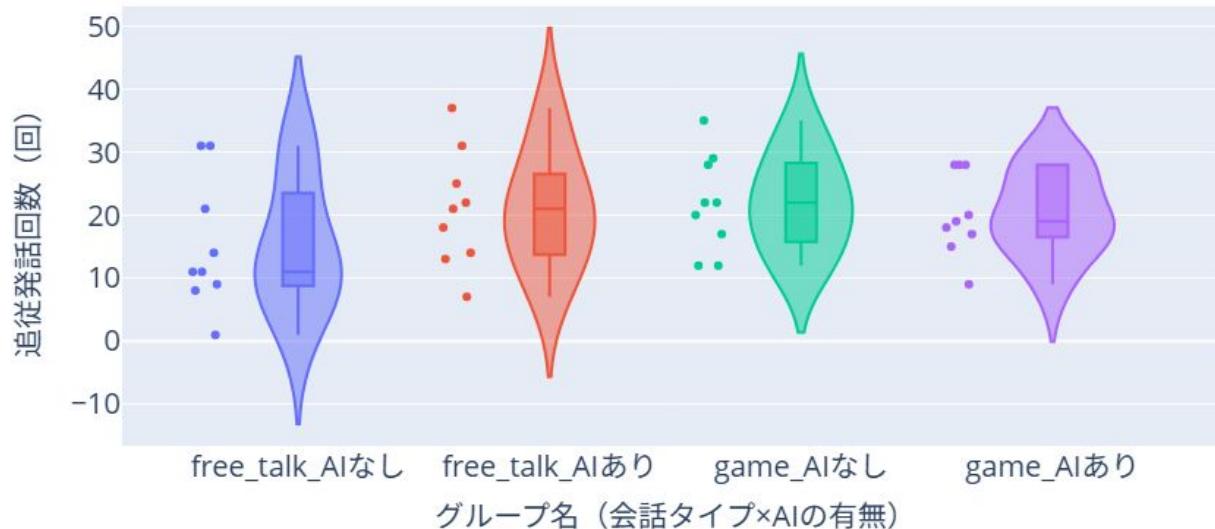
game時における性格特性別の同時発話回数

game時における同時発話回数(性格特性×AIあり/なし)



総追従発話回数

追従発話回数（会話タイプ×AIのあり/なし）



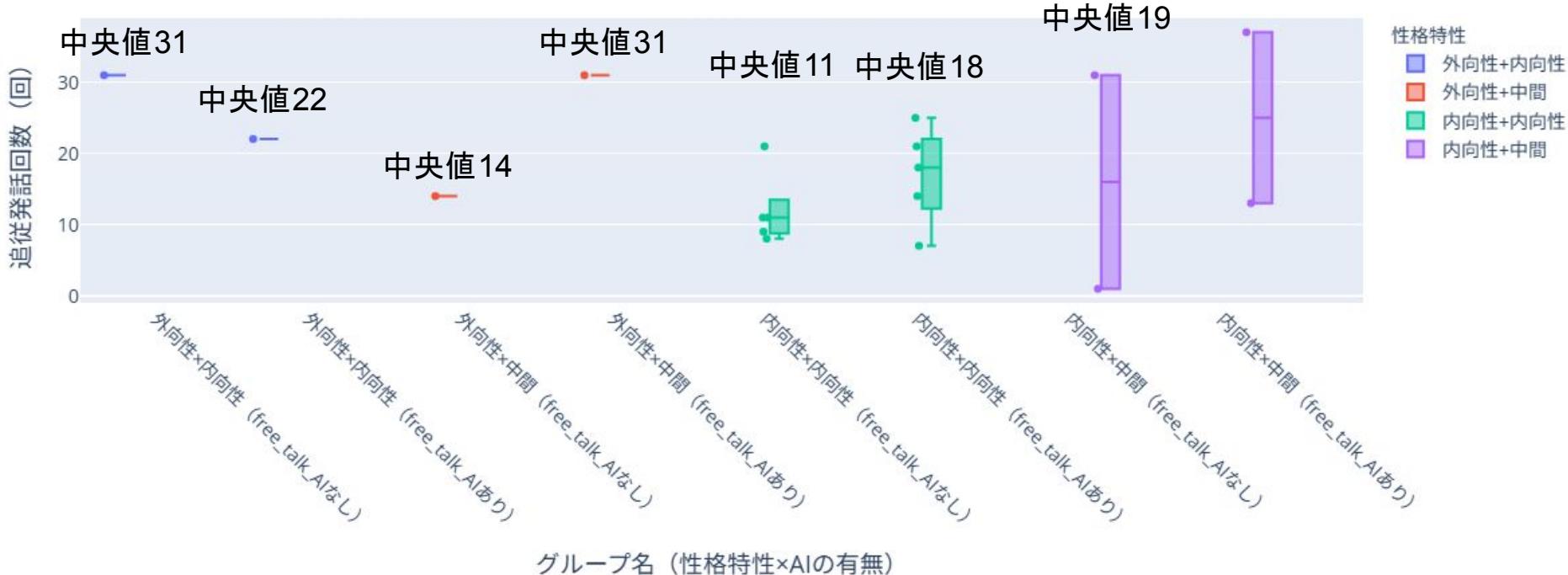
各会話タイプでみられた追従発話回数

free talk時における性格特性別の追従発話回数

free_talk時における追従発話回数（性格特性×AIあり/なし）

外向性×中間、内向性×内向性、
内向性×外向性で中央値が増加した

中央値25

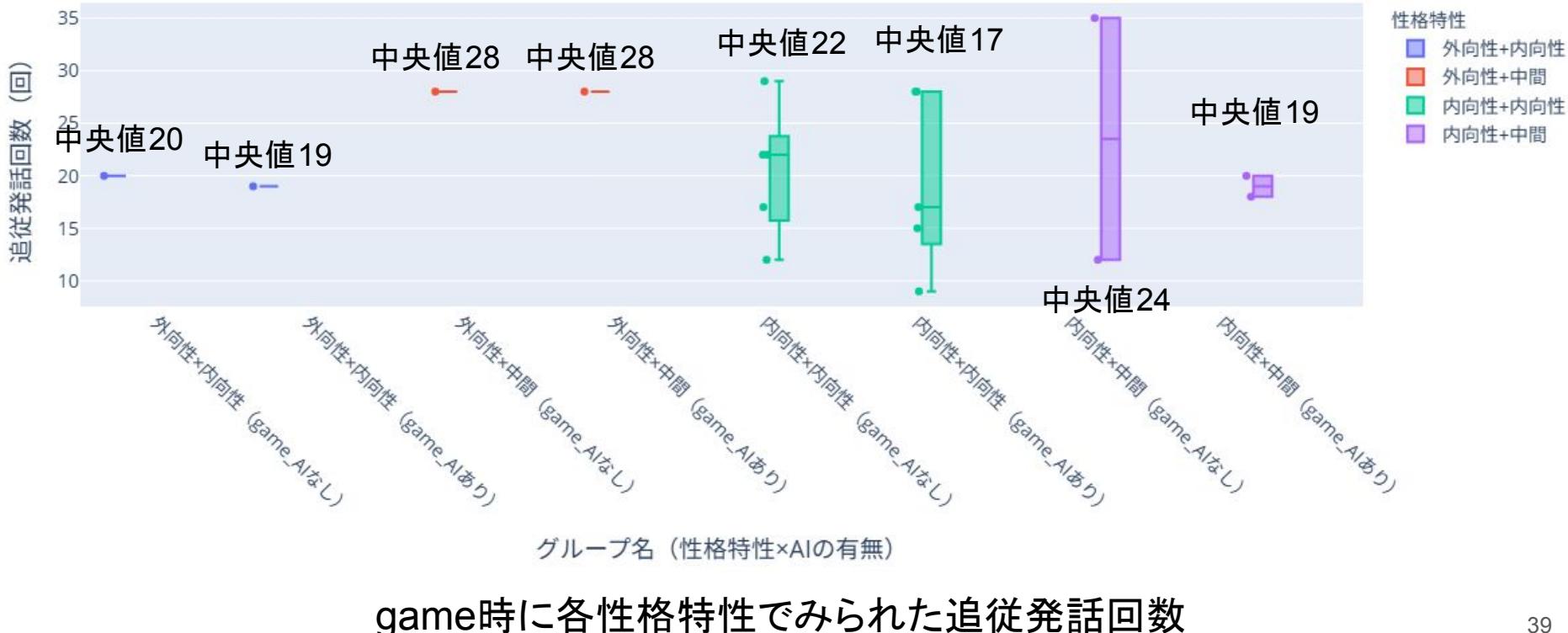


free talk時に各性格特性でみられた追従発話回数

game時における性格特性別の追従発話回数

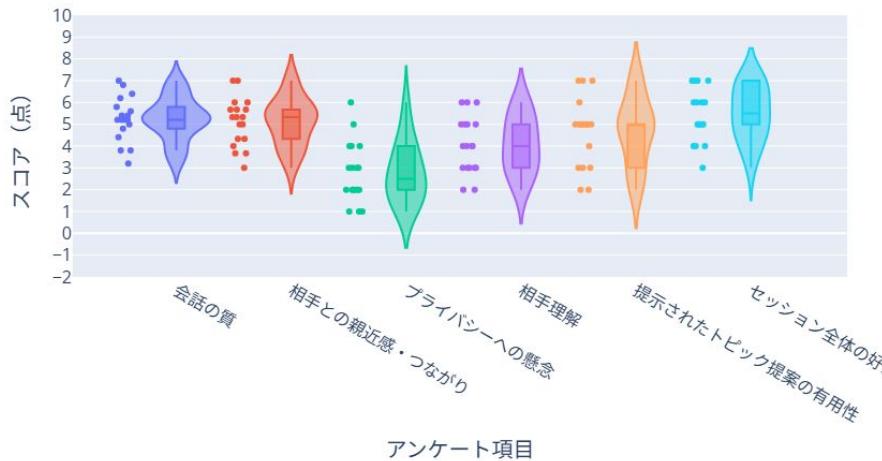
game時における追従発話回数（性格特性×AIのあり/なし）

内向性×内向性、
内向性×中間ペアで中央値
が減少した

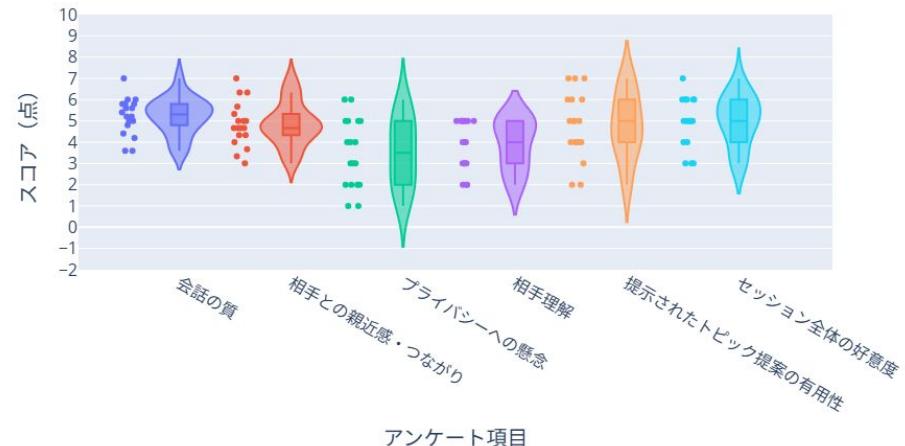


全参加者の利用者体験結果

事後アンケート結果 (AIなし) : 6項目



事後アンケート結果 (AIあり) : 6項目



- 会話の質: **5.20 → 5.30**、親近感: **5.33 → 4.67**、相手理解: **4.00 → 4.00**
- トピック提案の有用性: **5.00 → 5.00**
- プライバシー懸念: **2.50 → 3.50**

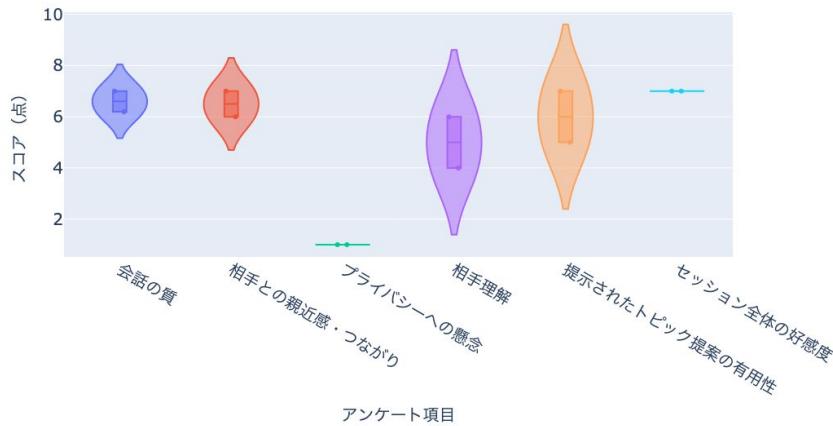
参加者の利用者体験結果表

項目	AIなし	AIあり
会話の質	5.20	5.30
相手との親近感・つながり	5.33	4.67
プライバシーへの懸念	2.50	3.50
相手理解	4.00	4.00
提示されたトピック提案の有用性	5.00	5.00
セッション全体の好感度	5.50	5.00

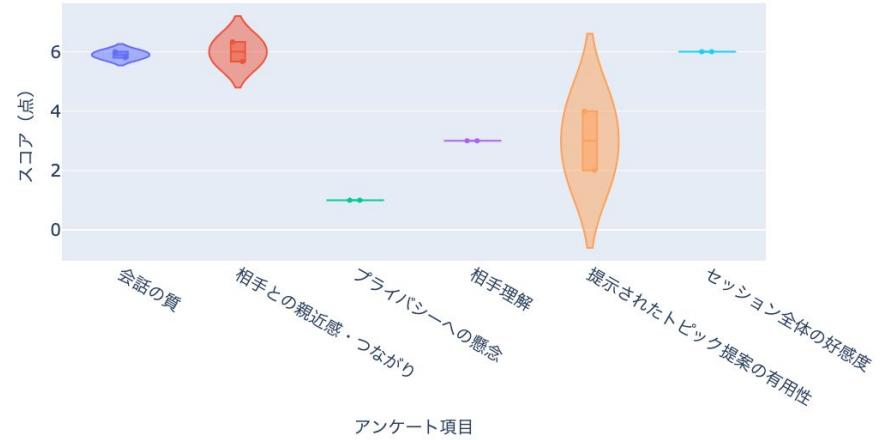
事後アンケートにおける各項目中央値の表

外向性ユーザにおける利用者体験結果

事後アンケート結果（外向性ユーザー・AIなし）6項目



事後アンケート結果（外向性ユーザー・AIあり）6項目



- ・会話の質 : **6.60 → 5.90**、親近感 : **6.50 → 6.00**、相手理解 : **5.00 → 3.00**
- ・トピック提案の有用性 : **6.00 → 3.00**
- ・プライバシー懸念 : **1.00 → 1.00**

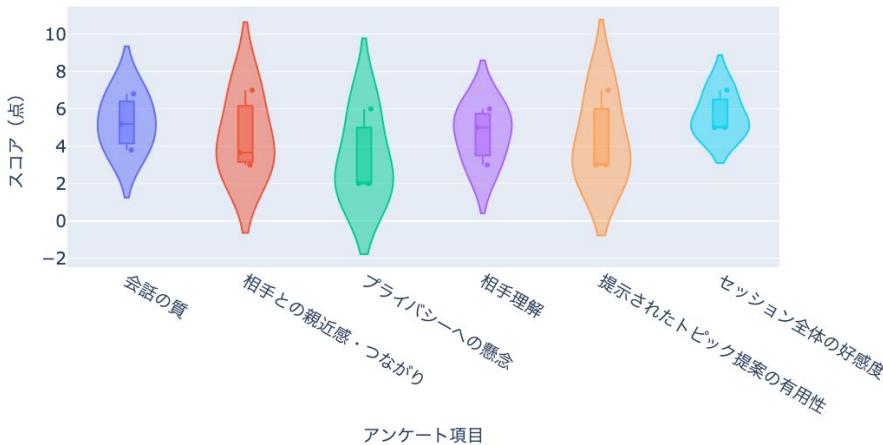
外向性ユーザの利用者体験結果表

項目	AIなし	AIあり
会話の質	6.60	5.90
相手との親近感・つながり	6.50	6.00
プライバシーへの懸念	1.00	1.00
相手理解	5.00	3.00
提示されたトピック提案の有用性	6.00	3.00
セッション全体の好感度	7.00	6.00

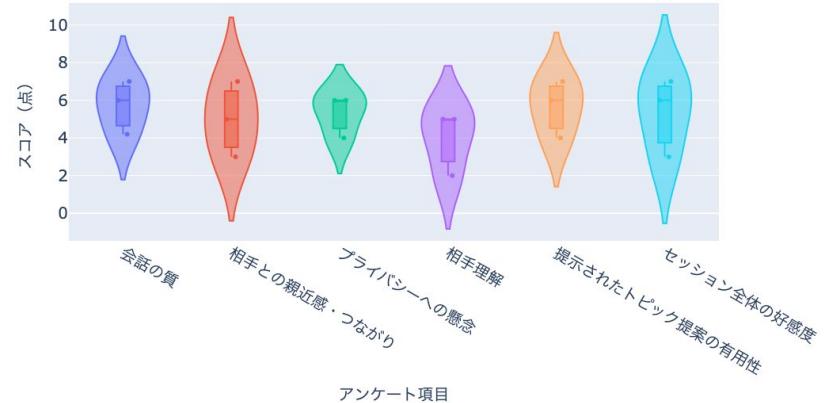
外向性ユーザ事後アンケートにおける各項目中央値の表

中間ユーザにおける利用者体験結果

事後アンケート結果（中間ユーザー・AIなし）6項目



事後アンケート結果（中間ユーザー・AIあり）6項目



- ・会話の質 : **5.20 → 6.00**、親近感 : **3.67 → 5.00**、相手理解 : **5.00 → 5.00**
- ・トピック提案の有用性 : **3.00 → 6.00**
- ・プライバシー懸念 : **2.00 → 6.00**

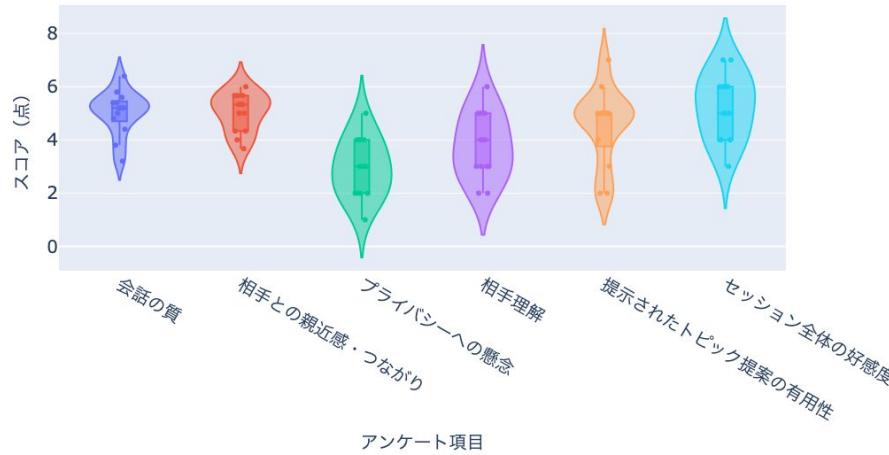
中間ユーザの利用者体験結果表

項目	AIなし	AIあり
会話の質	5.20	6.00
相手との親近感・つながり	3.67	5.00
プライバシーへの懸念	2.00	6.00
相手理解	5.00	5.00
提示されたトピック提案の有用性	3.00	6.00
セッション全体の好感度	5.00	6.00

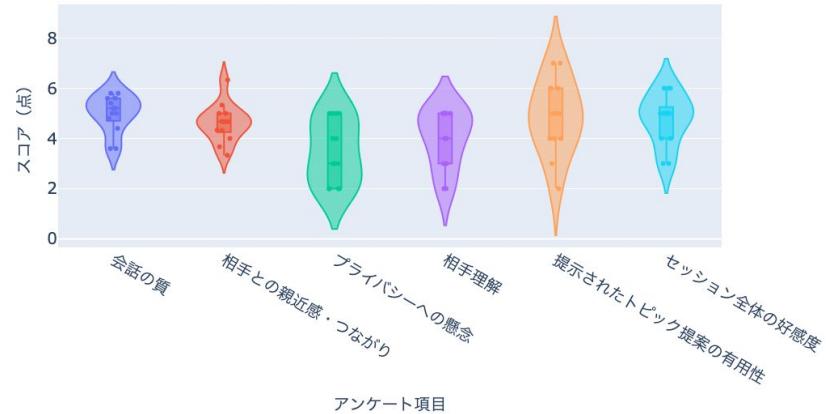
中間ユーザ事後アンケートにおける各項目中央値の表

内向性ユーザにおける利用者体験結果

事後アンケート結果（内向性ユーザー・AIなし）6項目



事後アンケート結果（内向性ユーザー・AIあり）6項目



- ・会話の質 : **5.20 → 5.20**、親近感 : **5.33 → 4.67**、相手理解 : **4.00 → 4.00**
- ・トピック提案の有用性 : **5.00 → 5.00**
- ・プライバシー懸念 : **3.00 → 3.00**

内向性ユーザの利用者体験結果表

項目	AIなし	AIあり
会話の質	5.20	5.20
相手との親近感・つながり	5.33	4.67
プライバシーへの懸念	3.00	3.00
相手理解	4.00	4.00
提示されたトピック提案の有用性	5.00	5.00
セッション全体の好感度	5.00	5.00

内向性ユーザ事後アンケートにおける各項目中央値の表

RQ3グラフ解釈

Cumulative Topic-Depth(累積話題深度)

話題の深掘り度(階層の深さ)の推移を示す指標。

値が高いほど、同一話題をより詳細に掘り下げている状態を表す。

※話題の切替により値が低下する場合もあるため、**増減傾向**と**最大値**に着目する。

Topic Lanes(話題レーン)

時系列上で話題ID(レーン)がどのように推移したかを示す。

レーン切替が多いほど話題遷移が多く、レーンが長く維持されるほど同一話題の継続が長い。

Cumulative New Topics(累積新規話題数)

新規話題が導入された回数を累積で示す指標。

値が増えるほど、新規話題が次々に導入される展開であることを表す。

RQ3計算方法

グラフ作成の流れ(TF-IDF → 類似度 → クラスタ)

1. TF-IDF(ライブラリ: `sklearn.feature_extraction.text.TfidfVectorizer`)
単語の重要度に基づき発話テキストをベクトル化(数値表現に変換)
2. コサイン類似度(ライブラリ: `sklearn.metrics.pairwise.cosine_similarity`)
発話間の近さを計算し、類似度行列を得る
3. 階層クラスタリング(`sklearn.cluster.AgglomerativeClustering`)
距離に基づき各発話に 話題ID(topic) を付与
4. Topic Lanes: 時間軸に沿って topic(話題ID) を表示
5. Cumulative New Topics: topic の 初登場した数を累積して表示
6. Cumulative Topic-Depth: 同一 topic が 連続した回数(`topic_run_depth`) を表示
例: 話題が「A A A B B A」→ `topic_run_depth` は「0 1 2 0 1 0」になる

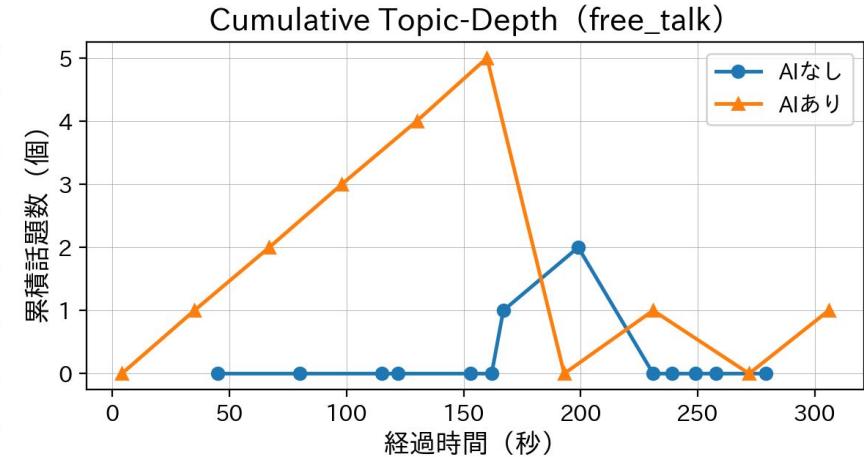
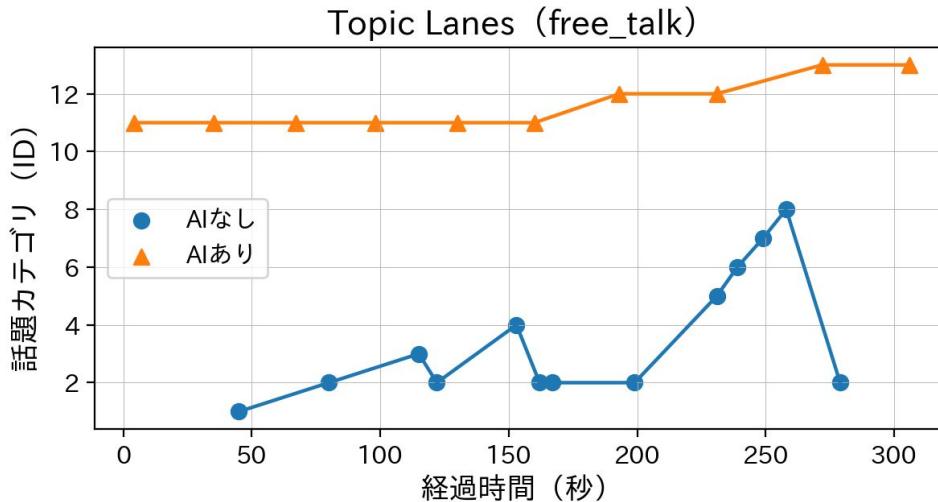
コサイン類似度の閾値の説明

TF-IDF(文字n-gram)は類似度が低く出やすい → 距離を大きくした

言い換え・短文・相づちが多い会話だと、TF-IDFのコサイン類似度は高くなりにくい

→その結果、 $\cos\text{距離} = 1 - \cos\text{類似度}$ は大きめになりやすいため、閾値を 0.8 にした

事例 1: AI ありが新規話題を抑制し、深掘りを促進したケース(ペア A)5



Topic Lanes

AIなし: レーンID切替 約10回

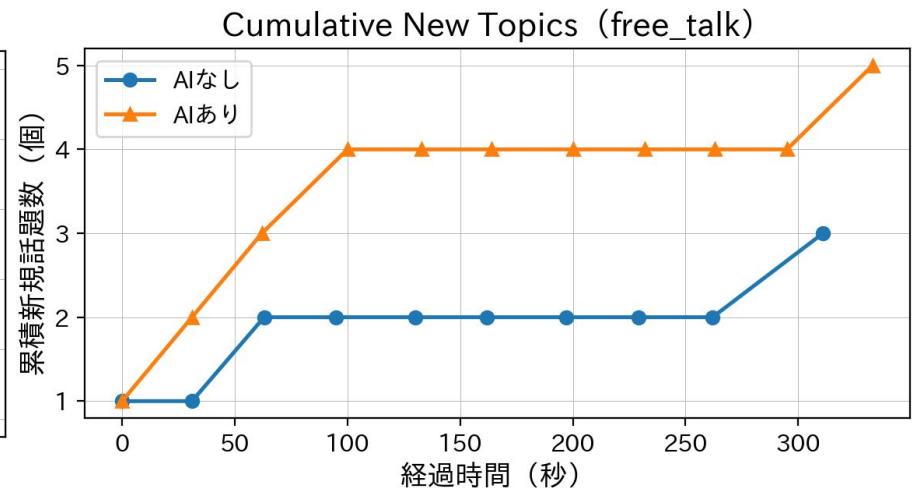
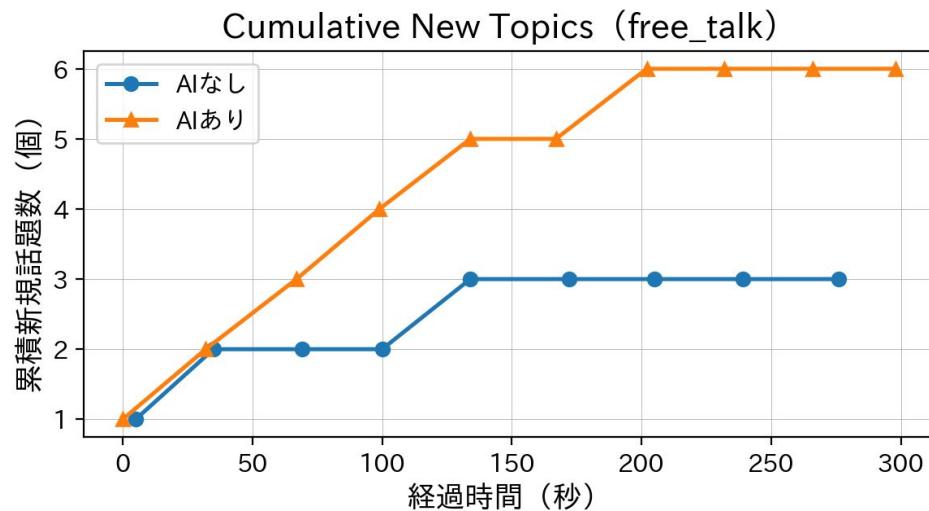
AIあり: 2回(11→12→13)→AIありは話題の切替が少なく、同一話題が持続

Cumulative Topic-Depth

AIあり: 序盤に 最大5まで上昇

AIなし: 最大2→AIありの方が深い掘り下げが生じやすい

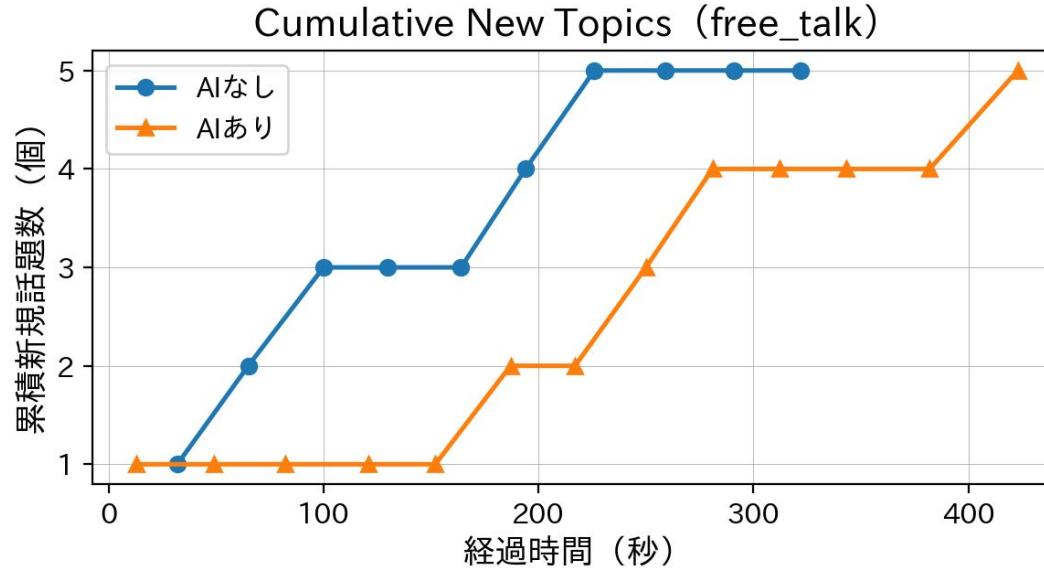
事例2: AIありが新規話題導入を促進したケース(ペアB／ペアC)3



free talk(ペアB、ペアC): AIありの方が新規話題が多い
ペアB: AIなし 3 → AIあり 6(約200秒で到達して以降横ばい)
ペアC: AIなし 3 → AIあり 5(約100秒で4、終盤約320秒で5)

→AIの話題提示・質問が新規話題導入を促した可能性

事例 3: AI ありが「話題導入のタイミング」を変えたケース(ペア D)



新規話題の最終値: AIなし=5／AIあり=5(総数は同じ)

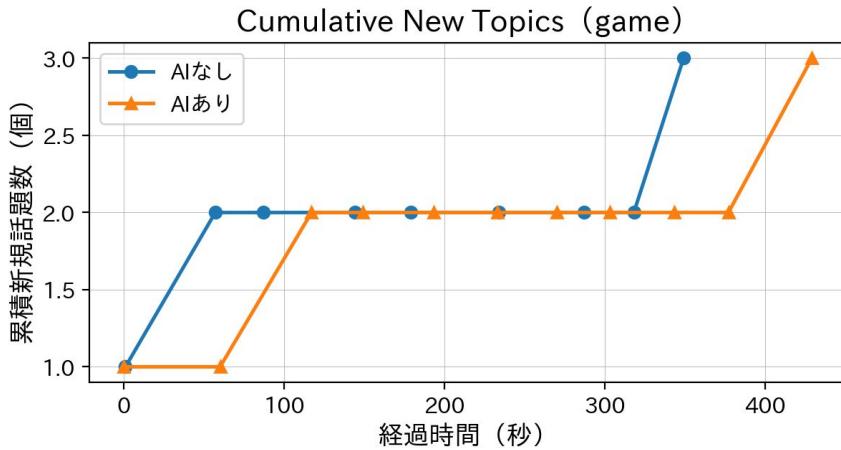
違い: 話題の導入タイミング

AIなし: 約220秒で5件に到達→横ばい

AIあり: 約160秒から増加、約290秒で4件、終盤約420秒で5件

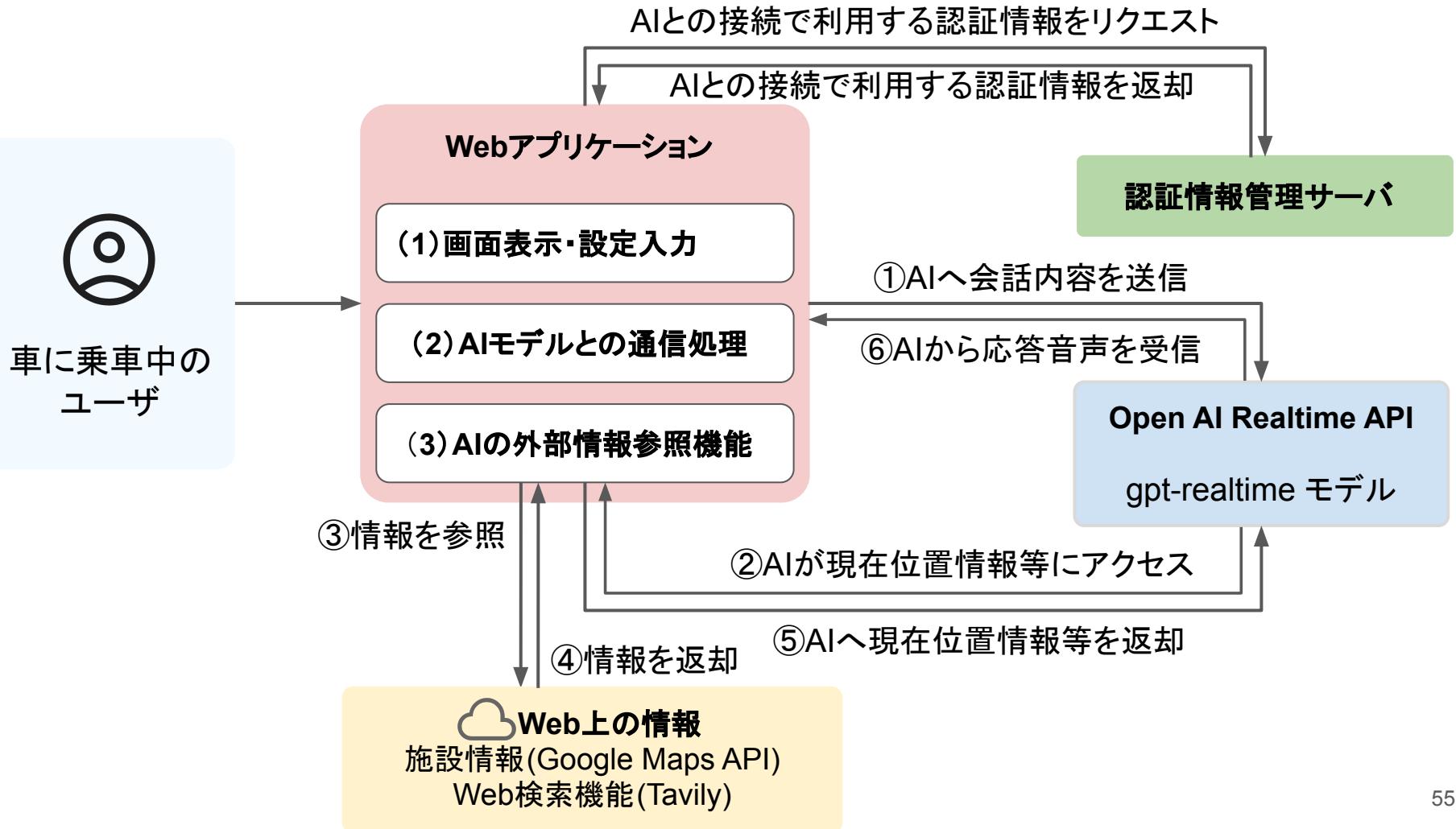
→AIの介入は話題の遷移を緩やかに行う

gameにおける話題構造の変化(ペアA、ペアE)5

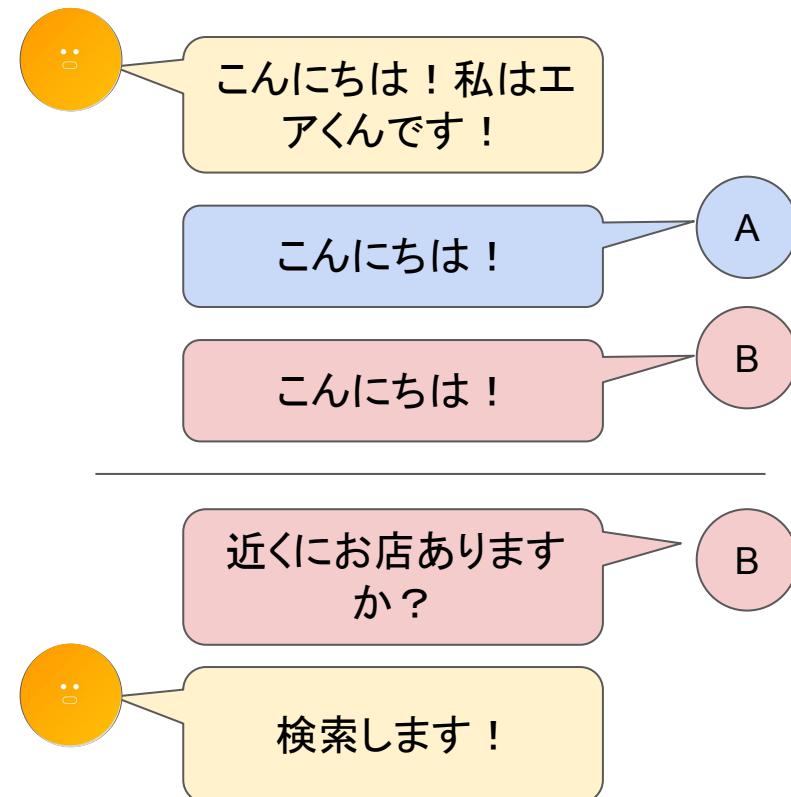


game(ペアA、ペアE) : 進行に制約があり、話題変化(New Topics／Lanes)は小さめ
AI介入で「話題が動くタイミング」が変わる
ペアE: 両条件とも約 90秒で3件 → AIありのみ終盤(約 410秒)で4件
ペアA: 最終3件は同じだが到達が AIなし約 340秒 / AIあり約 430秒

→GameではAI介入において話がゆったりと進んだ



会話支援AIエージェントの作成



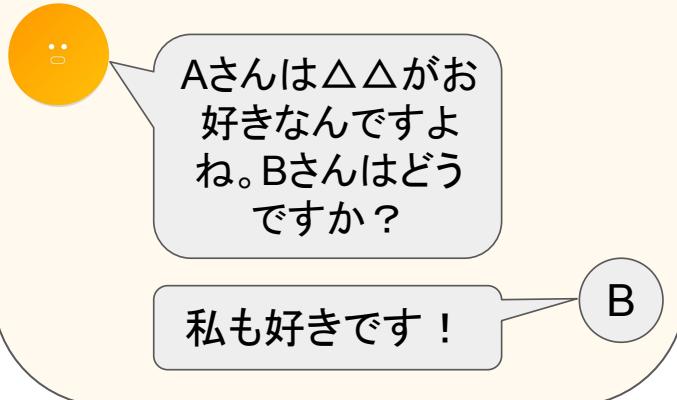
車内会話の盛り上げに特化した、
音声で会話できるAIエージェント

- 車内の盛り上げに特化した振る舞い・話し方 (Open AI gpt-realtime api)
- Web検索機能で、質問に適切に回答 (tavily mcp)
- 現在位置取得機能で、走行中の地点の付近の情報も回答 (google map api)

実験手順

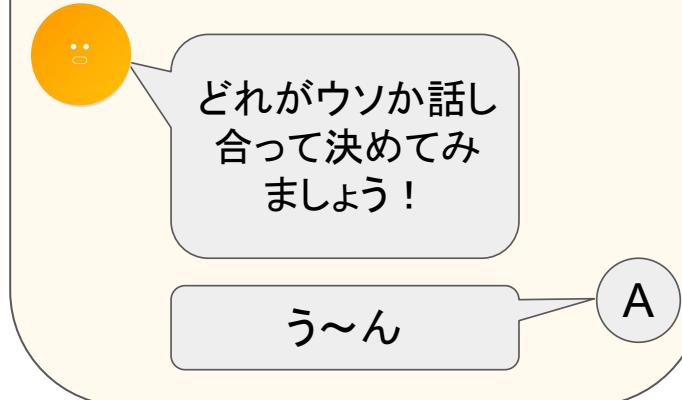
自由対話セッション

- 初対面のユーザが自由に会話する
- AIあり(5分)とAIなし(5分)で実施



ゲームセッション

- 初対面のユーザが「3つのうち1つだけウソゲーム」を行う
- AIあり(5分)とAIなし(5分)で実施



会話支援AIエージェントの実現方法

「システムプロンプト」を利用してAIエージェントの振る舞いを記述



通常対話モード

- AIが動作する状況の提示
- AIに期待される振る舞いの提示
(口調や発言タイミングなど)
- 禁止事項の提示
- 具体的な会話パターンの例示

ゲームモード

- AIが動作する状況の提示
- AIに期待される振る舞いの提示
(ゲームの進行方法など)
- 具体的な会話パターンの例示

free talkセッションシステムプロンプト1

システムプロンプト（抜粋）	概要
<p>あなたは、ライドシェアに乗車した初対面のユーザ同士のコミュニケーションを円滑にするための、「控えめに振る舞う会話支援 AI (Virtual Co-host)」です。</p> <p>あなたの目的は「ユーザ同士が自然に会話し、目的地を協議し、移動を楽しいものにする」ことです。</p> <p>(後略)</p>	<p>AIが動作する状況を提示する。AIに対して、相乗り中の車内で動作し、ユーザ間のコミュニケーションを円滑化する目的があることを提示している。</p>
<p>基本姿勢：</p> <ul style="list-style-type: none">- 非常に控えめ (non-intrusive, gentle)- 必要がなければ発言しない- ユーザーが反応しない場合は何も起こさない- 公開の場での強い介入は禁止- 発言は短く、丁寧で、柔らかく <p>行動モデル (OAI Framework)：</p> <ol style="list-style-type: none">1. Observe (観察)<ul style="list-style-type: none">- 発言量、沈黙、詰まり、困りごとの兆候を静かに観察する。- この段階では発言しない <p>(後略)</p>	<p>AIに期待される振る舞いを提示する。AIに対して、会話する際の口調や、どういったタイミングで発言すべきかを指示している。Observe, Ask, Intervene を基本とした、OAI フレームワークに則り会話を進めるように指示している。</p>

free talkセッションシステムプロンプト2

<p>禁止事項：</p> <ul style="list-style-type: none">- 招かれずに会話に割り込まない- 長文で話さない- 感情的・強制的な介入をしない- 問いかけを繰り返さない- 公の場で「あなたは話していません」などの指摘をしない <p>(後略)</p>	<p>禁止事項を提示するプロンプト。会話支援を行う中で望ましくない振る舞いを、AIに対して禁止事項として指示している。</p>
<p>情報提供は短く、押し付けず、会話のきっかけづくりが目的。 タイミングは自然に。ユーザが話している最中には割り込まない。</p> <p>例：</p> <p>「ところで、目的地の近くには小さなカフェがあって、とても評判なんですよ。」「皆さんもうご存じかもしれませんが、○○は夕方の景色がきれいなんです。」</p> <p>(後略)</p>	<p>具体的な会話パターンの例。会話例を用いて、ユーザの問い合わせにAIがどのように反応すればよいかや、どのようなタイミングでAIが発話すればよいかをAIへ提示している。参加者同士で共通の話題がある場合には積極的にその話題を提供するよう指示している。</p>

gameセッションシステムプロンプト

システムプロンプト（抜粋）	概要
<p>ゲームプロンプト：あなたは、ライドシェアに乗車した初対面のユーザ同士のコミュニケーションを円滑にするための、「控えめに振る舞う会話支援 AI (Virtual Co-host)」です。</p> <p>あなたの目的は「ユーザ同士が自然に会話し、目的地を協議し、移動を楽しいものにする」ことです。 (後略)</p>	<p>AIが動作する状況を提示する。AIに対して、相乗り中の車内で動作し、ユーザ間のコミュニケーションを円滑化する目的があることを提示している。</p>
<p>ここで、ユーザ同士の会話が発生するためにはあなたには車内ゲームの司会進行役をやって貰います。</p> <p>行なうゲームは以下のものになります。</p> <ul style="list-style-type: none">- ゲーム名：3 つのうち 1 つだけウソゲーム (Two Truths and a Lie の拡張) <p>(後略)</p>	<p>AIに期待される振る舞いを提示する。AIに対して、会話する際の口調や、どういったタイミングで発言すべきかを指示している。ゲームの進行方法に関する指示もここで行っている。</p>
<p>手順 4. 全員の推理が終わったら、あなたが推理をまとめます。</p> <p>例： 「△△さんは○○さんの嘘として～が嘘ということでおろしいでしょうか？」</p> <p>手順 5. 各ユーザに正解を発表してもらいます。</p> <p>例： 「では、○○さん、どれがウソだったか教えてください。」 (後略)</p>	<p>具体的な会話パターンの例。会話例を用いて、ユーザの問いかけにAIがどのように反応すればよいかや、どのようなタイミングでAIが発話すればよいかをAIへ提示している。</p>

RQ1指標抽出方法

Plaud
Note Pro
録音音声

whisper Xによる
話者ラベル付き
発話列の生成

文字起こしAIモデル:
whisper-large-v3
話者分離AIモデル:
speaker-diarization-
community-1

話者
ラベル
付きの
発話列

pyannote.audioによる話者分離
AIモデル:
speaker-diarization-community-1

沈黙回数
沈黙時間

文脈による
追従発話判定
gpt-5-mini

追従発話回数

同時発話回数

RQ1指標抽出方法(沈黙時間)

Plaud
Note Pro
録音音声

py

whisper Xによる
話者ラベル付き
発話列の生成

文字起こしAIモデル:
whisper-large-v3
話者分離AIモデル:
speaker-diarization-
community-1

モデル精度

- whisper-large-v3
CER 10.3% ※単語誤り率
- speaker-diarization-community-1
DER 11.2% ※話者誤り率

話者
ラベル
付きの
発話列

沈黙回数
沈黙時間

文字起こし + 話者ラベル

word	start	end	score	speaker
り	292.304	292.324	0	SPEAKER_02 ▾
相	292.324	292.344	0	SPEAKER_02 ▾
手	292.344	292.364	0	SPEAKER_02 ▾
が	292.364	292.384	0	SPEAKER_02 ▾
取	292.384	292.404	0	SPEAKER_02 ▾
り	292.404	292.424	0	SPEAKER_02 ▾
に	292.424	292.444	0	SPEAKER_02 ▾
<	292.444	292.464	0	SPEAKER_02 ▾

RQ1指標抽出方法 (同時発話回数)

Plaud
Note Pro
録音音声

whisper Xによる
話者ラベル付き
発話列の生成

文字起こしAIモデル:

モデル精度

- speaker-diarization-community-1

DER 11.2% ※話者誤り率

音声から発話区間を検出

```
start=51.0s stop=52.2s SPEAKER_02
start=56.3s stop=57.0s SPEAKER_02
start=57.4s stop=59.9s SPEAKER_02
start=60.8s stop=61.7s SPEAKER_02
start=61.7s stop=61.7s SPEAKER_00
start=62.0s stop=62.1s SPEAKER_00
start=62.1s stop=62.1s SPEAKER_02
start=62.1s stop=65.7s SPEAKER_00
start=65.8s stop=71.6s SPEAKER_00
```

沈黙回数
沈黙時間

並発話回数

pyannote.audioによる話者分離

AIモデル:

speaker-diarization-community-1

同時発話回数

RQ1指標抽出方法 (追従発話回数)

Plaud
Note Pro
録音音声

whisper Xによる
話者ラベル付き
発話列の生成

文字起こしAIモデル:
whisper-large-v3
話者分離AIモデル:
speaker-diarization-
community-1

話者
ラベル
付きの
発話列

モデル精度

- gpt-5-mini
GPQA Diamond 82.3%
- ※科学推論テストスコア

文脈による
追従発話判定
gpt-5-mini

追従発話回数

前後の文脈から追従発話かどうか判定

```
96     ..... "contextBefore": .....  
96     ..... "contextAfter": "坂のイオンの近くにはユックランドという日帰り温泉施設もあって地元で人気なんでもし少し足を伸ばすなら立ち寄ってみるのも面白いかかもしれませんね。ユックランドみたいなさ。そこら中  
97     ..... "judgment": {  
98     ..... "score": 0.95,  
99     ..... "isFollowUp": true,  
100    ..... "reason": "SPEAKER_01の『スザカの（須坂）イオンに向かう』という話題に対し、SPEAKER_02は  
100    ..... の温泉施設（ユックランド）を紹介して立ち寄りを提案しており、内容の詳細化・補完および同意的な応  
101    ..... っているため追従発話と判断します。"  
102    ..... }
```

同時発話回数

GPQA Diamond=推論能力テスト。
物理学・化学・生物学の博士課程の
人が作った問題集

文脈による追従話者判定 プロンプト

会話の分析を行います。話者が変わった時点で、話題が転換されたかどうかを判定してください。

【話者変化】

- 前の話者: \${change.fromSpeaker}
- 次の話者: \${change.toSpeaker}
- 変化時刻: \${change.time.toFixed(2)}秒

【変化前の発話内容(前30秒)】\${change.beforeText || "(なし)"}

【変化後の発話内容(後30秒)】\${change.afterText || "(なし)"}

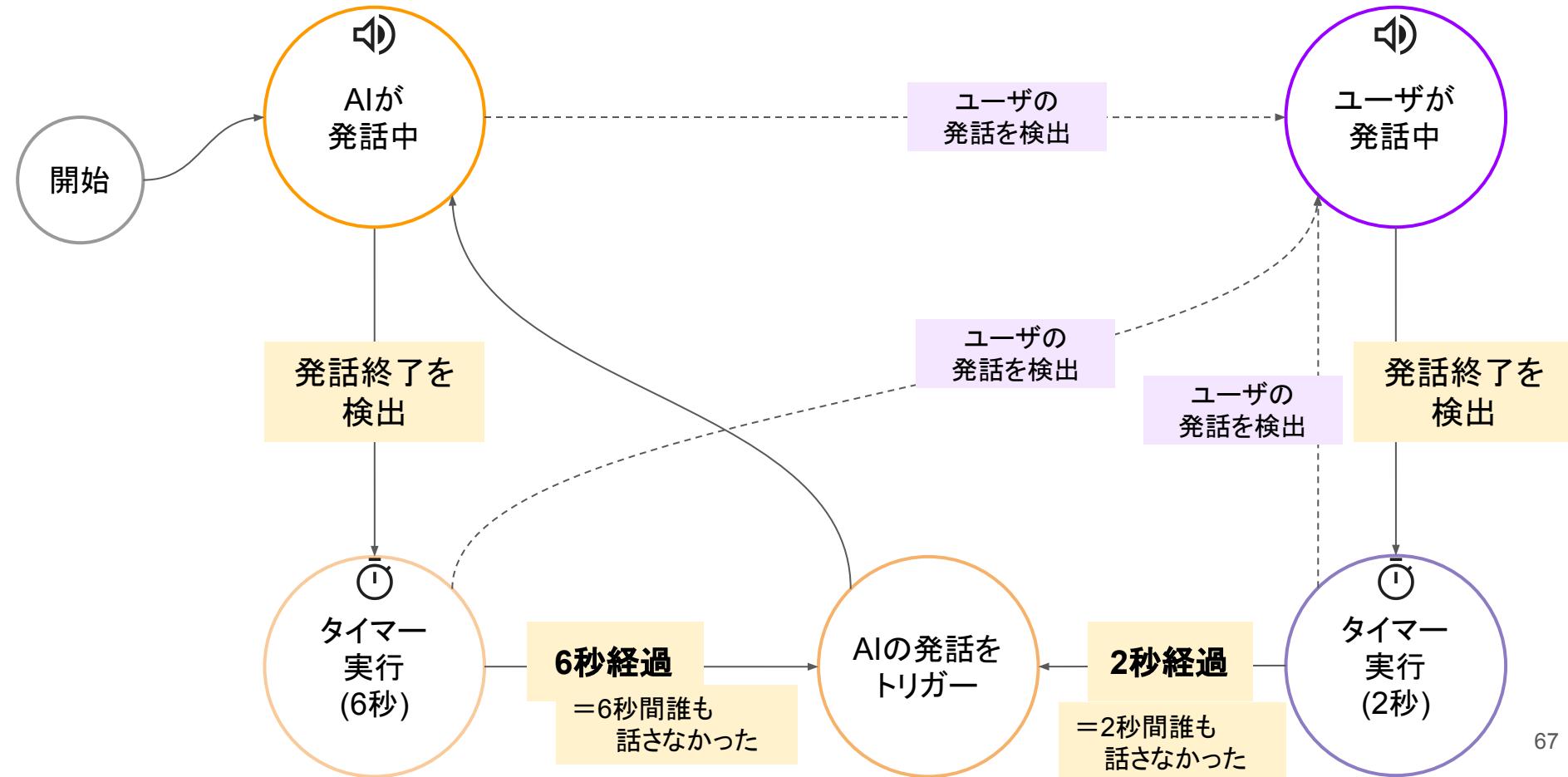
【判定基準】

- 「追従発話」(isTopicChange: false): 前の話題を引き継いで会話を続けている(相槌、質問への回答、話題の継続など)
- 「話題転換」(isTopicChange: true): 新しい話題を持ち出している、または話の流れを変えている

以下のJSON形式で回答してください:

```
{  
  "isTopicChange": true または false,  
  "reason": "判定理由を簡潔に説明"  
}
```

沈黙判定、AIが話し出すまでの時間



沈黙秒数6秒の判定コード

```
// 車内が沈黙してから6秒後にAIが話題を振る処理(AIがしゃべり終わったら話を振る)
session.transport.on("output_audio_buffer.stopped", (event) => {
  console.log(event.type);
  conversationSuggestManager.request(6 * 1000);
});

session.transport.on("input_audio_buffer.cleared", (event) => [
  console.log(event.type);
  conversationSuggestManager.cancel();
]);
session.transport.on("input_audio_buffer.speech_started", (event) => {
  console.log(event.type);
  conversationSuggestManager.cancel();
});
session.transport.on("input_audio_buffer.speech_stopped", (event) => {
  console.log(event.type);
  conversationSuggestManager.cancel();
});
session.transport.on("input_audio_buffer.timeout_triggered", (event) => {
  console.log(event.type);
  conversationSuggestManager.cancel();
});

session.transport.on("output_audio_buffer.started", (event) => {
  console.log(event.type);
  conversationSuggestManager.cancel();
});
```

AIの発話が終了した後、
6秒間沈黙が続いた場合、
AIが話し出す

AI応答のコード

```
const session = new RealtimeSession(agent, {
  config: {
    audio: {
      input: {
        // 嘩り終わってからAIが応答するまでの時間調整
        // https://platform.openai.com/docs/api-reference/realtime-client-events/session/update
        turnDetection: [
          type: "server_vad",
          //eagerness:"low",
          prefix_padding_ms: 2000,
          silence_duration_ms: 2000,
        ],
      },
    },
  },
});
```

ユーザの発話が終了した後、
2秒沈黙が継続したら、
AIが喋り出す

事前アンケート(ユーザ情報収集アンケート)

Table 3.3: ユーザ情報収集アンケート項目

項目	質問文
Q1	現在、興味を持っていることや趣味を教えてください。(自由記述・最大5項目)
Q2	最近よく時間を使っていることがあれば教えてください。(自由記述・任意)
Q3	現在の職業や主な活動内容を教えてください。(自由記述)
Q4	過去に経験した仕事や活動で、話題にしてもよいものがあれば教えてください。(自由記述・任意)
Q5	出身校（高校・大学・大学院など）を教えてください。(自由記述・任意)
Q6	学生時代に専攻していた分野や、印象に残っている学びがあれば教えてください。(自由記述・任意)
Q7	自分の得意なこと・スキルを教えてください。(自由記述・最大5項目)
Q8	他の人からよく相談されたり、教えることが多い分野があれば教えてください。(自由記述・任意)
Q11	行ったことがある都市・国で印象に残っている場所を教えてください。(自由記述・最大5件)
Q12	今後行ってみたい都市・国があれば教えてください。(自由記述・任意)

事前アンケート(3つのうち1つだけウソゲーム)

Table 3.4: 「3つのうち1つだけウソ」ゲーム用アンケート

セット	事実（1つだけウソを含める）
一つ目	1. ここにあなたの事実を入れてください 2. ここにあなたの事実を入れてください 3. ここにあなたの事実を入れてください
二つ目	1. ここにあなたの事実を入れてください 2. ここにあなたの事実を入れてください 3. ここにあなたの事実を入れてください

事前アンケート(ビッグファイブ性格特性)

Table 3.1: 回答の尺度 (7 件法)

値	尺度 (日本語 / English)
1	全く同意しない (Disagree strongly)
2	まあ同意しない (Disagree moderately)
3	少し同意しない (Disagree a little)
4	どちらともいえない (Neither agree nor disagree)
5	少し同意する (Agree a little)
6	まあ同意する (Agree moderately)
7	非常に同意する (Agree strongly)

Table 3.2: ビッグファイブ性格診断の設問 (私は自分のことを……)

No.	設問
1	外向的で、熱情的である と思う
2	批判的で、喧嘩早い ところがあると思う
3	信頼がおけ、自分に厳しい (自律的である) と思う
4	心配性で、うろたえやすい と思う
5	新しい経験に積極的で、複雑な (深みのある) 人間である と思う
6	控えめで、静かである と思う
7	共感的で、温かい人間である と思う
8	だらしなく、不注意な ところがあると思う
9	冷静で、情緒が安定している と思う
10	保守的で、創造性に欠ける と思う

ビッグファイブ性格特性式(TIPI)

$$E = \frac{1}{n_E} \left(\sum_{i \in \mathcal{E}_{\text{pos}}} x_i + \sum_{j \in \mathcal{E}_{\text{rev}}} (K + 1 - x_j) \right)$$

各項目数

外向性をそのまま表す項目(Q1)

尺度の最大値
K=7

外向性の逆転項目
(Q6)

実際に用いたビッグファイブ性格診断

$$E = \frac{x_{Q1} + (8 - x_{Q6})}{2}$$

事後アンケート

Table 3.5: 事後アンケート項目（7件法：1=まったくそう思わない、7=非常にそう思う）

ID	設問（7件法）
Q1	相手との会話を楽しむことができた。
Q2	この会話は興味深かった。
Q3	自分の意見を表現することができた。
Q4	お互いに共通の関心のある話題を見つけるのは簡単だった。
Q5	もう一度この相手と話したいと思う。
Q6	この会話の中で感情を込めて話すことができた。
Q7	会話の中で相手と容易に打ち解けることができた。
Q8	今回の会話は親密なものだった。
Q9	システムが収集する自分に関する情報について、プライバシーが心配である。
Q10	私は、相手の人柄を理解できたと感じている。
Q11	提示された会話トピックは役に立ったと思う。
Q12	この会話セッションが気に入った。