

複数エージェントによる文脈生成 AI を用いた情報共有システムの提案

Proposal for an Information Sharing System
Using Context-Generative AI with Multiple Agents笠原 有真¹⁾ 木村 泰知²⁾ 小野 哲雄¹⁾
Yuma Kasahara Yasutomo Kimura Tetsuo Ono

1 はじめに

デジタルサイネージが普及し、街頭や駅、商業施設、病院、大学など、あらゆる場所で目にすることができる [1]。デジタルサイネージは、従来の掲示板に比べてより視覚的な訴求力が高く、多くの情報を掲示することができる。また、コンテンツの変更も容易であるため、常に最新の情報を掲示することが可能である。

近年では、エージェントを用いたインタラクティブなサイネージも登場してきている。ユーザーの入力に応じて、コンテンツを変化させることができ、ユーザーがより必要とする情報を提示することが可能である。大学構内においても、学生に対しての情報の提示には掲示板やデジタルサイネージを用いて学生に情報を掲示することが多い。しかし、デジタルサイネージをはじめとした情報掲示システムは、ユーザーが能動的に確認しなければ情報を受け取ることができない。そのため、すべてのユーザーが情報を受け取ることができるとは限らない。

ところで、人間が情報を取得するのは掲示板やサイネージのみからではなく、周囲から聞こえてくる雑談から情報を取得するといった場合も多くある。例えば、以下のような会話を行っている人々とすれ違ったときを考えてみる。

- A 「履修登録ってもうした？」
- B 「まだしてない。いつが期限だった？」
- A 「来週とかじゃなかった？」

このとき、上記の会話を聞きつつすれ違った人は、それまで別のことを考えていた場合でも「履修登録の期限」を気にするといったことがある。つまり、人間は周辺で行われている何気ない雑談からも影響を受け、情報の取得や整理を行っていると考えられる。

そこで、本研究ではこの性質を利用し、ユーザーが日常の自然なコミュニケーションを通して必要な情報を受け取ることができるシステムを提案する。具体的には、生成型 AI により作成した文章を複数エージェントによって現実世界の文脈へ変換を行う。生成した文章を複数エージェント同士が音声による会話として出力することで、通りすがりに雑談としてユーザーへ影響を与える。さらに、会話によって生成された文脈へユーザーを引き込むことにより、エージェントによる情報共有を行うシステムの構築を目指す。

本稿では、複数エージェント同士の会話を生成するシステムおよび台本をもとに実際にエージェント同士の音声対話を行うシステムの構築を行い、実際の運用を通してユーザーの観察を行った結果について述べる。

2 関連研究

2.1 情報掲示システム

ユーザーにデジタルサイネージを注目させるための研究はいくつか行われている。小玉らは、歩行者に対してサイネージを正対表示させることにより、ユーザーの注目を獲得できることを明らかにした [2]。Alt らは、サイネージがインタラクティブであると、表示内容に対するユーザーの認知や記憶が向上することを明らかにした [3]。また、郷らは、インタラクティブなデジタルサイネージに対するユーザーの振る舞いを分析し、利用率を高めるための設計指針を提案した [4]。しかし、これらのシステムはユーザーの視覚へ訴え能動的に情報を取得することを期待しているため、ユーザーが受動的に情報を取得することは難しい。

2.2 エージェントを用いた情報掲示システム

ユーザーの注意を惹きつけるデジタルサイネージの形として、エージェントを用いたデジタルサイネージシステムの研究および実用化も行われている。三武らは、擬人化エージェントがユーザーに視線を向けることで、ユーザーの注意を惹くことを提案している [5]。また大浦らは、大学構内に音声対話可能なデジタルサイネージシステムを設置し、運用を行っていた [6]。2022 年に運用が終了したが、ユーザーからの注目度が高く、さまざまな場面で活用されていた。他にも、エージェントを利用したデジタルサイネージとして、株式会社ティファナ・ドットコムが開発している「AI さくらさん」¹⁾ が様々な場面で実際に運用されている。しかし、複数エージェントを用いたデジタルサイネージシステムについてはあまり運用されていない。

本研究では、ユーザーが受動的に情報を取得するために、複数のエージェントによる音声対話を用いる。音声を用いること、およびエージェントの対話を文脈に溶け込ませることでユーザーが通りすがりに受動的な情報取得を行う可能性を高める。また、会話の台本を生成型 AI によって作成することにより、低コストで複数エージェントの会話を実現する。

3 システム概要

本システムは、ChatGPT²⁾ を用いて雑談を台本として作成するシステムと、それにより実世界に文脈を生成するエージェントシステムに分けられる。

3.1 ChatGPT を用いた台本作成システム

会話は以下の三つの要素からなり、それぞれ ChatGPT を用いて生成を行った。

1) 北海道大学
2) 小樽商科大学

1) <https://www.tifana.ai/company>
2) <https://openai.com/blog/chatgpt>

小樽商科大学木村ゼミに所属するシーナ三姉妹の会話シミュレーションを行う。出力は以下のような形式で行うこと。
e.g. サイ: これは例です。

以下に制約条件を示す。

制約条件

- * 三姉妹は、長女のサイ・次女のエマ・三女のノイである。
- * サイは、真面目で優しくおしとやかだ。
- * サイは妹二人をそれぞれ「エマ」「ノイ」と呼ぶ。
- * サイは、「～です」「～ます」など、丁寧な口調。
- * サイはユーザーのことを「あなた」と呼ぶ。
- * エマは、優しくおしとやかで少し引っ込み思案だ。
- * エマは姉と妹をそれぞれ「サイお姉ちゃん」「ノイちゃん」と呼ぶ。
- * エマは、「～です」「～ます」など、丁寧な口調。
- * エマはユーザーのことを「あなた」と呼ぶ。
- * ノイは、明るく元気な少女で、中性的な言動を行う。
- * ノイは、シーナ三姉妹の末っ子である。
- * ノイは姉二人をそれぞれ「サイお姉ちゃん」「エマお姉ちゃん」と呼ぶ。
- * ノイは、「～だよ」「～よね」など、着飾らない口調。
- * ノイはユーザーのことを「君」と呼ぶ。

三姉妹が大学生活について雑談している。

図 1 台本作成時に入力する文章

以下の文章に対して、ユーザーが次にどのような文章を返すかを予測し、20文字以内で返してください。

サイ: みんな、大学生活はどう?
エマ: 私はまだ慣れてなくて、ちょっと疲れてるかな。
ノイ: 私は楽しいよ!でも、授業が多すぎる気がする。
サイ: そうだね、大学生活は充実してるけど、忙しい時もあるよね。

次に、ユーザーは以下のような発話を行った
ユーザー:

図 2 ユーザーの発話作成時に入力する文章

1. 台本
2. ユーザーの発話
3. ユーザーの発話を受けたエージェントの返答

まず、3体のエージェントの設定や口調を記述し、設定に準拠して特定の話題に関する台本を作成するように命令を与え、台本を作成する(1)。具体的には、まず三姉妹の会話シミュレーションを行うことを明示する文章を記述する。また、出力は「サイ: これは例です」という例を与え、話者とセリフをコロンで区切るように指定する。次に、3体がどのような関係性か、それぞれどのような性格か、語尾や話し方はどのようなものであるかを箇条書きで記述する。また、キャラクターごとのセリフ例をいくつか記述する。今回はそれぞれ3文ずつとした。最後に「三姉妹が〇〇について話している」といった文章を記述し、続きを ChatGPT に生成させることで三姉妹の雑談を作成した。実際に ChatGPT に入力した文章を図 1 に示す。出力は話者とセリフで区切り、プログラム上で扱いやすい形に整形後データベースへ保存した。

次に、(1)で作成した台本のどこまでを読み上げたタイミングでユーザーを会話に参加させるかを決め、ユーザーがインタラクションするための文章を生成する(2)、台本の文章をもとに、次にどのようなセリフが続くかを ChatGPT に出力させた。実際の文章を図 2 に示す。

最後に、ユーザーからのインタラクションを受けて、エージェントが発話する内容を生成する(3)。ChatGPT の API では、role というパラメータを渡すことで、出力の設定を行う文章であるか(system)、ユーザーからの入力であるか(user)、システムによって出力された文章で

あるか(assistant)を判別させることができる。そこで、(1)で記述した設定文章を system として設定し、それらに続くエージェント同士の会話を assistant として設定、(2)によって生成した文章を user と設定してそれらに続く文章を生成させることで、ユーザーがインタラクションを行った後の会話を作成した。

以上の文章は全てデータベース上にあらかじめ保存しておき、エージェントシステムから適宜取り出すようにした。

3.2 エージェントシステム

台本作成システムによって作られた台本を、デジタルサイネージに表示させた擬人化エージェントを用いて再生する。エージェントの-avatarとして、小樽商科大学木村ゼミナールのイメージキャラクターである、SEA-NA 三姉妹³⁾を用いた。理由としては、学内のキャラクターであることが認知されており、会話への引き込みが行いやすいと考えたからである、

エージェントの表示には、55 インチの縦型ディスプレイを3台用いた。エージェントに対してインタラクションを行うためのインターフェースとして、キーボードを中央のディスプレイ前に配置した。ディスプレイ及びキーボードは図 3 のように配置した。

会話を行う際の処理の流れについて記述する。まず、台本作成システムから、台本とそれに紐づくユーザーのインタラクション用文章を取得する。取得した台本をもとに、エージェントのセリフについては VoiceVox⁴⁾を用いて台本の音声合成を行う。その際、異なるエージェント同士が会話していることをわかりやすくするため、エージェント毎に音声を変更した。また、エージェントが発話中かどうかをわかりやすくするために、発話中と待機中でモーションを切り替えた。指定されたタイミングで、ユーザーのインタラクション用文章を吹き出しとして表示させる。図 4 の上の吹き出しが画面上に表示されたタイミングでユーザーがキーボードを押すと、図 4 の下の吹き出しに変わり、インタラクションしたことを視覚的に知ることができる。キーボードが押されたタイミングで、エージェントが続きの返答の発話を行う。ユーザーからのインタラクションが無かった場合は、そのまま台本の最後まで会話を続ける。各会話は、話題ごとに5秒のインターバルを設け、切り替わりがわかりやすいようにした。

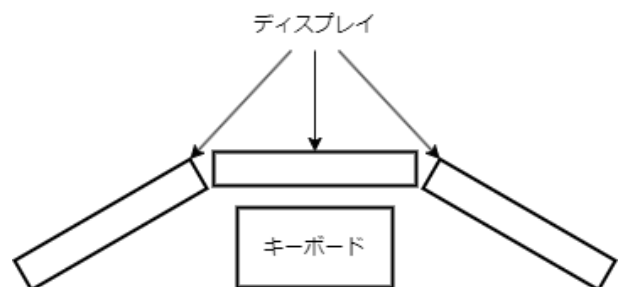


図 3 ディスプレイおよびキーボードの配置

3) http://www.sea-na.net/?page_id=3241

4) <https://voicevox.hiroshiba.jp/>

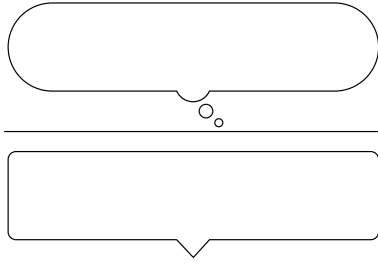
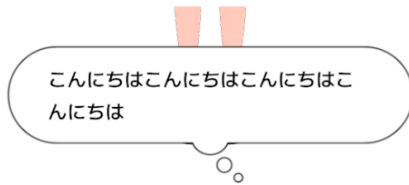


図 4 吹き出しの種類 (上がインタラクション前, 下がインタラクション後)

実験中



上のような吹き出しが出ているとき、
Enterキーを押すとアバターが反応します
(何もしなくてもアバターは話し続けます)

図 5 インタラクション可能であることを示す貼り紙

4 実験

4.1 実験手法

本システムがユーザーの引き込みを行うことができるかを検証するために大学構内における実験を行った。システムの設置場所として、多くの学生が通る図書館前の通路を選択した。学生が多く通る時間帯として、11時50分から12時50分までの1時間稼働させ、その間にサイネージ前を通過したユーザーの反応を観察した。エージェントの会話内容としては、通過するユーザーに身近な話題とするため、大学生活に関する雑談や、大学のお知らせといった内容を用いた。サイネージへのインタラクションが可能であることを示すため、キーボードの手前に図5のような貼り紙を掲示した。実際の実験の様子を図6に示す。

4.2 実験結果

実験中にサイネージ前を通過したユーザーの反応を、(1) 通過のみ、(2) 視線を向ける、(3) インタラクションの三つに分類しそれぞれの人数を計測した結果を表1に示す。

表 1 実験結果	
通過のみ	129 人
視線を向ける	140 人
インタラクション	24 人

実験中にサイネージ前を通ったユーザーはは 293 人だった。その内、視線を向けて興味を示したのが 140 人、キーボードを押してエージェントの会話に参加したのは 24 人であった。また、エージェントの会話に 1 人



図 6 実験の様子

で参加したのは 2 人であった。2 人以上で参加したのは 22 人であった。

5 考察

本実験では、サイネージの前を通ったユーザーのうち、56%のユーザーが興味を示していた。このことから、ある程度ユーザーの注意を惹くことができたと言える。しかし、サイネージ前を再度通過した際に注意を向けないユーザーもいた。このことから、目新しさが注意を引く大きな要因であった可能性が高いため、今後実験条件を整えて検証を行う必要がある。

注意を引くことはできた一方、エージェントの会話へ参加したのは 8%のみであったため、文脈への引き込みは十分とは言えない。また、会話に参加はしたが、その後最後までエージェントの会話を聴いていないユーザーもいた。そこで、会話に参加したユーザーの行動を分析し、それぞれ何が要因であったのか考察を行う。

5.1 会話を最後まで聴いていたユーザー

ユーザーがエージェントの会話を最後まで聞いていたときの行動について、実際の例を挙げて分析を行う。そのとき再生していたエージェントの台本と、ユーザーの一連の行動を図7に示す。図中の文章で、[]が使われている箇所はシステムの動作を表し、()が使われている箇所はユーザーの行動を表している。まず、ユーザーは会話の内容を聴いていたわけではないが、システムを見てディスプレイに近づいた。そのとき、タイピングとしては次のエージェントの発話と重なっていたためユーザーへの反応が遅れたが、ユーザーはエージェントからの反応を待機していた。これは、挟まったエージェントの発話である、「いいね、その時は私がおごるからね。」という文章が短かったからではないかと考えられる。また、ユーザーが送った文章を受けて会話を続けていることがエージェントの発話からわかりやすかったことも、ユーザーが最後まで会話を聴いていた要因の一つであると考えられる。

サイ：みんな、大学生活は楽しんでる？
 エマ：はい、私は大学の雰囲気が好きです。でも、まだ友達が少なくって…
 ノイ：私は大学生活がとても楽しいよ！友達もたくさんできたし、新しいことにチャレンジするのも好きだから。
 サイ：エマは引っ込み思案だから、友達ができるのに時間がかかるかもしれないけど、焦らずに自分のペースで頑張っってね。
 [「応援してるよ！頑張っってね。」という吹き出しが表示される]
 エマ：はい、頑張ります…サイお姉ちゃん。
 ノイ：私もエマお姉ちゃんのこと応援するよ！一緒に遊んで、リラックスして、友達ができるように手伝ってあげるよ。
 サイ：そうだね、大学生活は勉強だけではなく、友達との時間も大切だからね。
 ノイ：あなたは大学生活を楽しんでる？
 サイ：はい、とても楽しんでるよ。でも、勉強もしっかりしないとね。
 エマ：そうだね、大学は自分自身を成長させるための場所だから、勉強も頑張らないと。
 ノイ：私は勉強も楽しいけど、やっぱり友達との時間が一番かな。
 サイ：それはそれでいいと思うよ。大学生活はそれぞれの楽しみ方があるからね。
 エマ：そうですね。私も自分らしく楽しみたいと思います。
 (ディスプレイに近づき、貼り紙を見る)
 ノイ：私も！今度、あなたたちと遊びに行きたいな。
 (キーボードを操作)
 サイ：いいね、その時は私がおごるからね。
 ユーザー：応援してるよ！頑張っってね。
 サイ：あなたに応援されると、私たちが頑張れますね。ありがとうございます。
 (再度キーボードを押す)
 (少し離れる)
 エマ：そうだね、私たちは大学生活でたくさん学んで成長できるように頑張りたいね。
 ノイ：うん、でも私たちが頑張れば頑張るほど、将来の選択肢も増えてくるよな。

図 7 ユーザーが最後まで会話を聴いていたときの台本と行動

5.2 途中で会話から離れたユーザー

次に、ユーザーが最後まで会話を聴いていなかったときの行動について分析を行う。本実験を通して、会話への参加後に最後までエージェントの会話を聴いていなかったときのユーザーの行動として、立ち去る前にキーボードを連打するといったことが多くみられた。実際の例を図 8 に示す。例示したように、一度キーボードを操作して吹き出しが変化したあともキーボードを連打する行動が見られた。このことから、キーボードを押した後の変化が発話内容のみであり、エージェントの振る舞いは変化しないため、エージェントの会話に参加した実感があまり湧かなかったことが引き込みができていない要因の一つと考えられる。また、今回の例ではユーザーを指し示す単語(「あなた」など)がエージェントの発話に出てきていないため、発話内容から変化が分かりづらかったことも考えられる。そのため、インタラクションを行った際に、視線や身体動作といったエージェントの振る舞いにも変化を加えることで、引き込みを継続させることが可能になるのではないかと考える。

6 まとめ

本稿では、複数エージェントを用いて、ユーザーへの情報共有を行うシステムの提案を行った。具体的には、生成型 AI を用いて作成した文章を複数エージェント同士の会話として実世界の文脈へと埋め込み、その文脈へユーザーを引き込むことを目的としたシステムを構築した。実際に大学内にて稼働させてユーザーの観察を行った。結果として、注意を引くことはできたが、ユーザーを引き込むというところについては課題が残った。目新しさが注意を引いた要因となっている可能性があるため、長期的な運用を行って効果検証を行う必要がある。

今後は、よりユーザーを会話へと引き込むことができるようなインタラクションデザインを考える必要がある。本システムではユーザーへのフィードバックが不十

サイ：こんにちは。あなたたち、今日は元気ですか？
 エマ：はい、お姉ちゃん。元気です。あなたはいかがですか？
 ノイ：おはよう、君たち！元気だよな？
 サイ：ええ、私も元気です。最近、何か面白いことがあった？
 [「最近は特に面白いことはなかったです。」という吹き出しが表示される]
 エマ：あのね、私最近、お料理を始めたの。お姉ちゃん、食べに来てくれない？
 ノイ：私も食べたいな！お姉ちゃん、私も行くよ！
 サイ：そうだね、エマちゃんの料理は美味しいから、ぜひ食べに行こう。
 (ディスプレイに近づき、貼り紙を読む)
 (キーボードを押す)
 (ディスプレイを見る)
 ユーザー：最近は特に面白いことはなかったです。
 (ディスプレイを見ながらキーボードを連打する)
 サイ：そうですか、退屈な毎日は嫌ですね。私は最近、読書にはまっています。
 (その場を離れる)
 エマ：あら、サイお姉ちゃんはいつも偉いわね。私は最近、お菓子作りにハマっています。
 ノイ：それは良いね、私は最近友達と一緒に野球をやっているんだ。風が気持ちいいし、体を動かすのは楽しいよ。

図 8 ユーザーが途中で離れたときの台本と行動

分であり、会話の引き込みがあまり行われなかった。そのため、どのようなエージェントの動きや反応のタイミングが会話に参加した感覚を生み出すのかを検証し実装することで、よりユーザーの情報取得を支援することが可能になるのではないかと考えられる。

参考文献

- [1] 中村伊知哉. デジタルサイネージの動向. 情報管理, Vol. 55, No. 12, pp. 891-898, 2013.
- [2] 駿小玉, 翔太須藤, 進洪沢. デジタルサイネージに向けた情報を正対表示させ歩行者を引き付けるシステム. Technical Report 4, 茨城大学大学院理工学研究科, 茨城大学大学院理工学研究科, 茨城大学工学部, mar 2014.
- [3] Florian Alt, Stefan Schneegass, Michael Girgis, and Albrecht Schmidt. Cognitive effects of interactive public display applications. In *Proceedings of the 2nd ACM International Symposium on Pervasive Displays*, PerDis '13, p. 13-18, New York, NY, USA, 2013. Association for Computing Machinery.
- [4] 郷健太郎, 石丸尚輝, 飯塚重善, 木下雄一朗. インタラクティブ・デジタルサイネージに対する気づきの調査. ヒューマンインタフェース学会論文誌, Vol. 17, No. 4, pp. 363-372, 2015.
- [5] 裕玄三武, Wu Hsuehhan, 晶一長谷川. キャラクタを用いたデジタルサイネージが通行人の注意を引きつけるための視線制御. エンタテインメントコンピューティングシンポジウム 2018 論文集, 第 2018 巻, pp. 276-281, sep 2018.
- [6] 大浦圭一郎, 山本大介, 内匠逸, 李晃伸, 徳田恵一. キャンパスの公共空間におけるユーザー参加型双方向音声案内デジタルサイネージシステム (<特集> 音声対話システムの実用化に向けて). 人工知能, Vol. 28, No. 1, pp. 60-67, 2013.