

ユーザ生成型音声対話システムにおけるクリエイターとユーザの相互刺激によるインセンティブ向上の検討

飯塚 遼¹ 李 晃伸¹

概要：対話の内容から対話エージェントのモデルや動き等を含めた音声対話システムコンポーネントを「音声対話コンテンツ」として、自由に作成・編集・共有するユーザ生成型音声対話システムの成立に向けて、コンテンツ作成・登録・利用に対するユーザの自発的なインセンティブを強化する仕組みを検討する。システム利用者を利用形態から閲覧中心のユーザとコンテンツを作成・登録するクリエイターに分類した上で、それぞれの内部、あるいは相互のインセンティブ刺激を、コンテンツ統計情報の提示やユーザ間発話共有によって促す仕組みを提案する。本発表では、MMDAgent を用いて実証実験を行った結果を報告する。

1. はじめに

音声対話システムは、近年の研究によって実環境で運用される機会も多くなり、その有用性が明らかになった [1]。しかし、実環境で広く利用されているとは言えないのが音声対話システム利用の現状である。

ところで、Wikipedia[2]、YouTube[3]などの消費者主導型メディアが近年注目を集めている。この消費者主導型メディアでは、利用者がコンテンツによるサービスをただ享受するだけでなく、利用者が自らコンテンツの作成や評価、改良を行っている。利用者が作成したコンテンツのことをユーザ生成型コンテンツという。管理者・利用者ともにユーザ生成型コンテンツを利用する最大の利点は、利用者の要望をそのままリアルタイムにシステムへ反映することができる点や多彩なコンテンツを生み出すことができるという点である。また、利用者がコンテンツ作成を行う理由としては、活動自体に楽しさを感じる [4] や不完全な情報を補足したい [5]、自己表現を行いたい [6] などが挙げられる。こうした理由から、作成者にとって金銭的な見返りがなくてもコンテンツ作成が行われ、他の閲覧する利用者にとっても有益なコンテンツが作成される。

こうしたユーザ生成型コンテンツの利点に注目し、過去の研究において、音声対話システムにユーザ生成型コンテンツの機能を取り入れた、ユーザ生成型音声対話システムが考案された [7]。このシステムでは、システムの対話内容を一般利用者が作成できるのが特徴である。これにより、

これまでの管理者によって登録された対話内容しか利用者は享受できないという受動的なシステム利用から、最新の話題や利用者の欲しがる情報を利用者自身によって提供するという能動的なシステム利用が可能となる。その結果、対話内容の登録が増加することで対話コンテンツが充実し、利用者が惹きつけられるような音声対話システムが実現すると期待できる。

先行研究 [7] では、ユーザ生成型音声対話システムを公共空間に約 1 ヶ月間設置し、利用状況を観察した。その結果、利用者の自発的で多様な登録が見られた。このことから、音声対話システムにおいてもユーザ生成型コンテンツ成立の可能性が示唆された。

そこで、本研究では利用者間で相互に刺激しあうような仕組みをシステムに導入し、ユーザ生成型音声対話システム実現のための成立要件を探る。

2. ユーザ生成型音声対話システム

2.1 音声対話コンテンツとは

音声対話システムは、声や動き対話シナリオなどのタスク依存の部分と、音声認識エンジンや音声合成の方式などのタスク非依存の部分に分けることができる。そこで、本研究ではタスク依存の部分音声対話コンテンツ、タスク非依存の部分処理モジュールとして、コンテンツとシステムの分離を考える。

2つの対応関係を図 1 に示す。音声対話コンテンツには音響モデルや言語モデル、声質モデルなどの各モデルと、単語辞書、対話シナリオ、モーション、エージェントモデルがある。処理モジュールには、音声認識部、対話管理部、音声合成部、エージェント描画部がある [8]。

¹ 名古屋工業大学大学院 工学研究科
Graduate School of Engineering, Nagoya Institute of Technology.

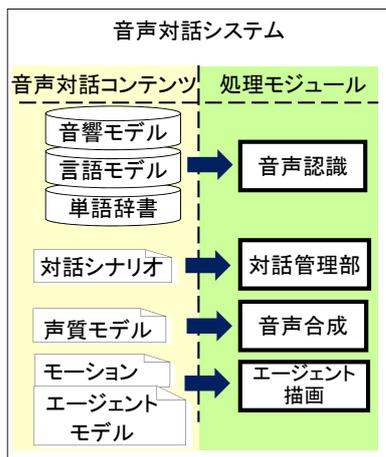


図 1 各音声対話コンテンツと各モジュールの対応関係

2.2 ユーザ生成型音声対話システムとは

ユーザ生成型音声対話システムは、ユーザが自在にコンテンツを作成し、共有し評価しあうユーザ生成型メディアとして成立するような基盤となる音声対話システムである。前節で示したようなコンテンツの分離が行われることで、ユーザによるあらゆる音声対話コンテンツの自由な作成が可能となる。さらに、コンテンツを作りやすくしたり、共有する仕組みを確立することで、最終的にはユーザ生成型メディアへの展開を期待するものである。

我々は現在、ユーザの記述する音声対話コンテンツの範囲を認識キーワードと応答文の組に絞り、それらを自由に登録できる音声対話システムを対象として、ユーザ生成型音声対話システムの成立要件について研究している。

2.3 ユーザ生成型音声対話コンテンツ成立のための課題

先行研究 [9] において、認識キーワードと応答文を誰でも登録可能なユーザ生成型音声対話システムを構築し、キャンパス内の公共空間に 20 日間設置して利用状況を観測した。その結果、一定の利用者数があったものの自発的にコンテンツを登録する利用者はごく少数であった。利用者動向の分析から、以下のような課題が挙げられた。

- 応答文登録者と音声対話システム利用者の相互刺激
音声対話システムの利用を増やすためには、応答文を登録する利用者と音声対話システムを利用したい利用者が互いに影響しあうような関係を築き、相乗効果を生み出すような仕組みを音声対話システムに導入することが重要になる。しかし、現状のユーザ生成型音声対話システムにはこうした相互刺激を促す仕組みが実装されていない。
- システムへの話し掛けの促進
音声対話システムは、実社会において普及しているとは言い難く、利用者の多くはシステムに向かって声を掛けること自体に抵抗感を持っている考えられる。そうした現状では、話し掛けに対する熱意が失われシ

テムが使われなくなってしまうと考えられる。

- コンテンツ登録に対するインセンティブの強化
コンテンツ作成環境を整備することで、コンテンツをつくる楽しさを増幅させることができる可能性がある。

3. クリエイターとユーザの相互刺激によるインセンティブ向上

本研究では、利用者集団をユーザとクリエイターに分離して、互いに刺激しあうような仕組みを導入することで、ユーザ生成型音声対話システムの成立を目指す。

3.1 クリエイターとユーザ

先行研究 [9] では、利用者がシステムに慣れ親しむことが、ユーザをコンテンツ作成へと誘引するとの前提で研究を行い、利用者に Felica を利用してユーザ登録をしてもらい、システム利用の回数に応じて応答を変えることで、繰り返し使わせ、コンテンツ作成へ引きこもうとした。一方で、コンテンツ作成にはクリエイター登録が必要であり、ユーザ登録とは独立した登録が必要であった。

実験結果として、132 名のユーザ登録が得られ、16 名のクリエイター登録があった。しかし、クリエイター登録者のうち、実際にコンテンツ作成をしたのは 5 名で、そのうちの 4 名はユーザ登録をしていなかった。この 4 名はシステムへの親しみとは別の動機でコンテンツを作成したと考えられる。ユーザ登録を行ってからコンテンツ作成を行った利用者はほとんどいなかった。

このことから、コンテンツ生成が中心の利用者とシステムへの話し掛け中心の利用者は性質が異なることがわかった。そこで、本研究では利用者集合を 2 つに分割し、コンテンツ生成中心の利用者をクリエイター、話し掛け中心の利用者をユーザと定義し、それらごとのインセンティブ*1を議論することとする。

3.2 ユーザ・クリエイターの相互刺激によるインセンティブ付与

利用者集団としてユーザとクリエイターの行動原理が異なるため、インセンティブもそれぞれを個別に考慮する必要があると考えられる。そうすることで、ユーザ・クリエイター間で相互にインセンティブを与え合いシステム利用の促進が期待できる。ここでは、ユーザ生成型メディア成立の要となる利用者間でのインセンティブの相互付与に着目し、具体的に以下の 3 つの相互刺激関係を考える。

- ユーザ間での相互刺激
共有発話履歴を表示することでユーザ間での相互刺激を促進する。これは他の人も話しているという状況を提示することでユーザ発話の誘発を狙う。

*1 使ってみたい、やってみようという行動刺激

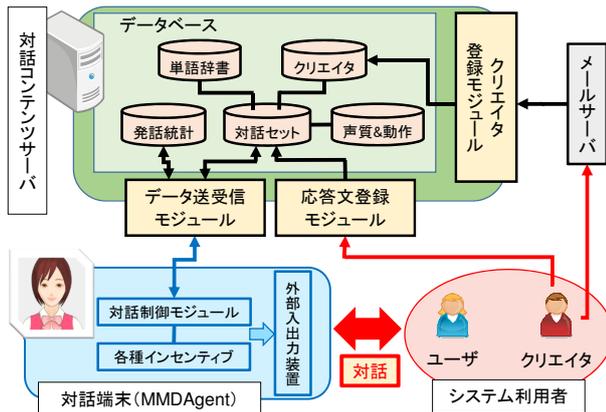


図 2 提案システムの構成図



図 3 提案システムのスクリーンショット

● ユーザ・クリエイター間での相互刺激

発話数によるランキング・新着コンテンツ提示を行うことでユーザ・クリエイター間での相互刺激を促進する。クリエイターは、ランキングから自身の作成したコンテンツのユーザ評価を獲得でき、コンテンツ作成への意欲向上につながると期待できる。ユーザはランキングによる話題のコンテンツわかることでユーザ間での話題共有ができ、新着コンテンツの提示では新規コンテンツの利用ができる。

● クリエイター間での相互刺激

クリエイター間での相互刺激としても、ランキング・新着コンテンツ提示が挙げられる。ランキング提示から、クリエイター間での競争が期待できる。また、新着コンテンツ提示から他のクリエイターの動向がわかることで、コンテンツ作成の手助けになる。

4. 提案システム

前節で述べた提案を導入したユーザ生成型音声対話システム(以下提案システム)を開発した。提案システムは本研究室でこれまで運用されてきたユーザ生成型音声対話システムを一部ベースにし、本研究室で開発された音声インタラクションシステム構築ツールキット MMDAgent[10]を用いて設計したシステムである。このシステムは、音声対話コンテンツをサーバからダウンロードして実行するため、ネットワークにつながっている Windows PC とマイク、スピーカの 3 点があれば利用可能である。

提案システムの構成図を図 2、スクリーンショットを図 3 に示し、以下対話コンテンツサーバと対話端末、認識キーワードと応答文(以下対話セット)の登録について述べる。

4.1 対話コンテンツサーバ

対話コンテンツサーバの役割を以下に記述する。

- 対話端末とのデータ送受信と各種情報の保持
端末への対話シナリオの送信と端末から送られてくる

認識結果の受信を行う。

● 各種情報の保持

各クリエイターの応答登録数やユーザの発話数、取得した認識結果から各対話セットの累積応答数を保持する。他にも、単語辞書や声質・モーションモデルが保持されている。

● ウェブページの提供

対話セットを登録するためのウェブページをクリエイターに提供する。

● クリエイター情報の保持

クリエイターのメールアドレスと、それに紐付けられたパスワードを保持する。

4.2 対話端末

対話端末上では、まずユーザ・クリエイター間の相互刺激として、図 3 左側に示しているように総合ランキングと新着キーワードを表示する。総合ランキングは対話セットが利用された回数によって生成し上位 7 個を表示する。新着コンテンツは、対話セットが登録された時間と対話シナリオの更新時間との差分を併せて、新規対話セットを 7 個表示する。また、ユーザ間の相互刺激として、サーバを介した全ユーザの共有発話履歴をリアルタイムに表示する(図 3 右側)。

システム処理の流れを説明する。MMDAgent に各機能をプラグインとして実装した。対話端末を起動すると対話端末はデータ送受信モジュールを通して、応答文データと単語辞書を内部に保存する。ユーザが接続されたマイクに発話を行うと、対話制御モジュールにおいて、音声認識結果と対話セットを比較し、対話シナリオ内に認識結果とマッチするものがあれば、その応答文を音声合成しユーザへ出力する。このときの発話選択の結果は対話コンテンツサーバにも保持され、その回数もカウントされ、回数は総合ランキングに使われる。さらにユーザの発話で応答を返



図 4 対話セット登録のトップ画面

せなかったものがあれば、未登録キーワードとして、対話コンテンツサーバ内に保持するように設計した。

4.3 対話セットの登録

ユーザは Web ページから任意のコンテンツ (対話セット) を登録できる。

図 4 に対話セット登録のウェブページのトップ画面を示す。対話セット登録画面では統計情報を見せることで、クリエイター間の相互刺激を高める。中央の表には、そのユーザがこれまでに登録した対話セットを表示する。右下の薄赤色の表は対話セット登録者のランキングである。左側が対話セット登録数が多い人、右側が登録した対話セットの累計応答数が最も多い人の順にランキングしてある。

図 5 に対話セットの登録画面を示す。認識キーワード、応答文、声質、動作を入力できる。さらに、図 5 の右側には応答数が多かった発話のランキング (人気キーワード) と、未登録キーワードのランキングを提供し、クリエイターへの応答文登録の補助を行う。未登録キーワードは平仮名 1 文字の認識結果は対象外として定義した。

コンテンツ検索画面 (図 6) では、キーワードを入力することで登録されている応答文データベース内のコンテンツを検索できる。キーワードとの部分一致またはクリエイター毎に検索可能である。

5. 評価実験

5.1 実験条件

提案システムに実装した相互刺激を行う仕組みが機能するかを調べるために、実環境で運用し利用状況の集計を行った。実験条件を表 1 に示す。当初調査期間を 1 月 10 ~ 17 日としたが期間後にも利用があったため、23 日までを調査期間とした。

被験者は音声対話システム (MMDAgent) 利用方法の学内講習を受けた学生で、少なくとも一つの対話セットを登録する条件で自由に利用するよう研究室単位で依頼した。

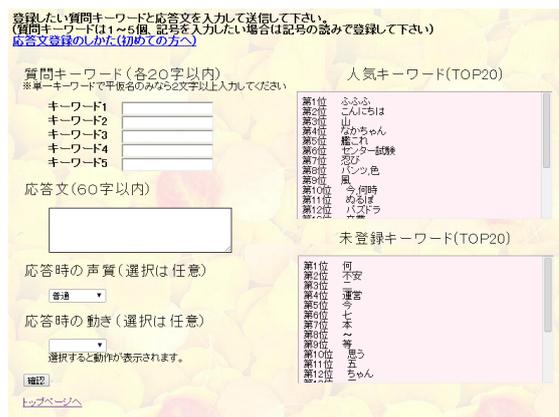


図 5 対話セット登録画面



図 6 対話セット検索画面

表 1 実験条件

調査期間	2014 年 1 月 10 日 ~ 2014 年 1 月 23 日
登録者	大学生/大学院生 (23 人) 学科内他研究室
実験環境	各個人の Windows PC
初期コンテンツ数	40
取得データ	被験者数, コンテンツ数, 発話数, 応答数, アンケート

そのため、被験者はユーザでありクリエイターでもある。実験は各被験者の個人 Windows PC で行った。初期コンテンツとして、大学周辺施設の案内、挨拶、雑談など 40 個の対話セットを初期登録した。

被験者には各項目 5 段階評価 (5 が最もアンケートの項目内容に当てはまる) のアンケートに回答してもらう。各項目毎のラベルとアンケートの内容を表 2 ~ 4 に示す。

5.2 実験結果

表 5 に調査期間中に取得したデータを示す。なお、ユーザがシステムに話しかけた回数をユーザ発話数 (入力は雑音トリガを含む)、ユーザの話し掛けに対してシステムが応答を返せた回数を応答成功数とした。調査期間中に実験に参加した被験者は 23 人、実際にコンテンツ登録を行ったのは 19 人であった。その中でアンケートに回答したのは 19 人であった (コンテンツ登録者とは一致しない)。各データの遷移の様子を図 7 および 8 に示す。

表 2 ユーザ間の相互刺激

ラベル	アンケート内容
A1	発話履歴の表示でシステムに話し掛けやすくなったと感じたか
A2	発話履歴の表示で応答文登録をしてみようと感じたか

表 3 ユーザ・クリエイター間とクリエイター間の相互刺激

ラベル	アンケート内容
B1	コンテンツ順位がわかることは応答文作成の動機向上になったか
B2	総合ランキングはシステムへの話し掛けの手助けになったか
B3	総合ランキングを見ることで応答文登録をしてみようと感じたか
B4	新着コンテンツはシステムへの話し掛けの手助けになったと感じたか
B5	新着コンテンツを見ることで応答文登録をしてみようと感じたか
B6	人気キーワードの表示は応答文作成に役立ったか
B7	未登録キーワードの表示は応答文作成に役立ったか

表 4 総合評価

ラベル	アンケート内容
C1	今後もこの音声対話システムを利用しようと思うか
C2	今後のこのシステムへの期待 (5 が最も期待できる)
C3	総合評価 (5 が最も良い評価)

表 5 日ごとに得られたデータの数

日付	被験者数 累計	新規対話セット 登録数	ユーザ 発話数	応答 成功数
10-13日	0	0	0	0
14日	6	10	492	70
15日	8	8	409	64
16日	12	7	148	30
17日	20	29	1708	212
18-19日	20	0	0	0
20日	21	4	30	2
21日	21	0	0	0
22日	23	2	90	28
23日	23	3	109	11
合計	23	63	2986	417

図 7 から、被験者数が増えると対話セットも増えていることがわかり、被験者が増えていない日には対話セットも増えていないことがわかる。このことから、今回の実験では被験者は実験に初めて協力した日に、対話セットの登録をしやすという傾向があることがわかった。同様にユーザ発話数も被験者が増えた日に増加していることから、今回の実験ではクリエイター・ユーザ共にシステムの繰り返し利用をさせることが少なかったと言える。図 7 および 8 で、17日に最もデータが集まったのは事前に告知した終了日(17日)を意識した利用者によって、提案システムが使用されたためであると考えられる。

5.3 グループ分け

対話セット登録数をもとに被験者を便宜的にユーザとクリエイターの 2 カテゴリーに分類した。図 9 に対話セット登録数に基づくグループ分けの内訳を示す。今回の実験では、被験者に少なくとも一つ対話セット登録を依頼しており、

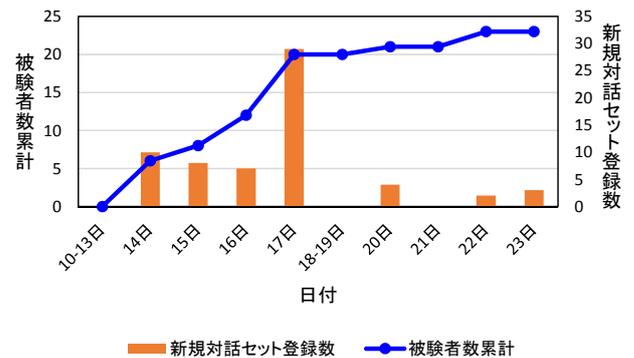


図 7 被験者数累計と対話セット登録数の遷移

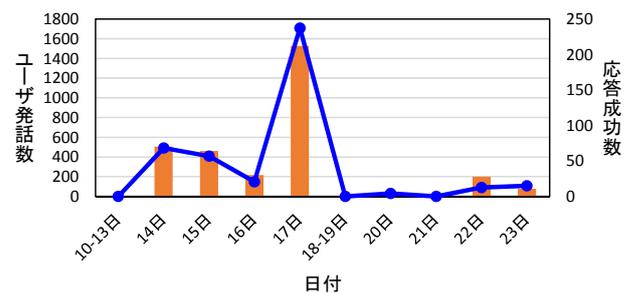


図 8 ユーザ数発話数と応答成功数の遷移

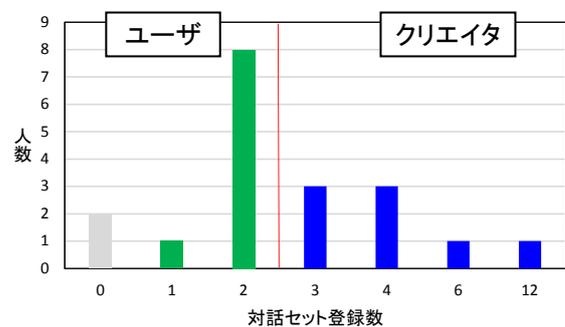


図 9 グループ分けの内訳

個人のモチベーションによる対話セット登録ではなかったということ、登録数を観察したときに登録数 2 の分布が最も多かったことから、登録数が 3 以上をクリエイター、2 以下をユーザとしてグループ分けを行った。なお評価において、対話セットを登録していない被験者のアンケート結果は用いていない。

5.4 ユーザ間の相互刺激に関する考察

表 6 にユーザ間の相互刺激に関するアンケート結果を示す。全体的に発話履歴は話しかけやコンテンツ登録に関して有効に機能することがわかった。また、この効果が大きく現れたのはユーザで、ユーザのクリエイター化やシステムへの話しかけを促進する要因になったと考えられる。

表 6 ユーザ間に関する相互刺激

		A1	A2
クリエイター	平均	3.5	2.8
	分散	1.5	1.9
ユーザ	平均	3.8	3.4
	分散	0.8	1.6

表 7 ユーザ・クリエイター間とクリエイター間に関する相互刺激

		B1	B2	B3	B4	B5	B6	B7
クリエイター	平均	4.3	4.5	3.9	4.0	3.8	4.1	2.5
	分散	1.7	1.0	1.4	2.3	1.9	1.6	1.0
ユーザ	平均	3.9	4.3	3.4	4.2	3.6	4.1	3.0
	分散	1.0	1.1	0.9	0.8	1.1	0.5	1.8

表 8 総合評価に関するアンケート結果

		C1	C2	C3
クリエイター	平均	3.9	4.4	4.5
	分散	1.4	0.5	0.3
ユーザ	平均	3.7	4.4	4.1
	分散	1.1	0.2	0.3

5.5 ユーザ・クリエイター間とクリエイター間での相互刺激に関する考察

表 7 にユーザ・クリエイター間とクリエイター間の相互刺激に関するアンケート結果を示す。話しかけに関して (B2, B4) は大変高い効果があったことがわかる。これはニコニコ動画 [11] と同様にランキング化したコンテンツ (対話セット) を提供することは、ユーザ・クリエイター間やクリエイター間に対して利用刺激を与えることができるということを裏付ける結果となり、相互刺激は適切に機能したと言える。

また、システムへの発話時に収集した統計情報をランキング化し、対話セット登録時に提示することは登録を促進する大きな要素になったと言える (B1, B6, B7)。未登録キーワードは高い評価を得ることが出来なかったが、これは今回の未登録キーワードの条件では、対話セット生成に役立たない漢字一文字などのキーワードが多くなってしまったことが原因と考えられる。

5.6 提案システムの総合評価に関する考察

表 8 は、提案システムに対する総合評価である。提案システムは全体的に被験者に好印象を持たれた。特に、今後の期待 (C2) が高いことから、提案した相互刺激によりユーザ生成型音声対話システムが今後ユーザ生成型メディアとして発展する可能性が示唆されたと言える。

6. むすび

本研究では、ユーザとクリエイターの間での相互刺激の促進することで、ユーザ生成型音声対話コンテンツが成立する条件を探った。そのために対話コンテンツのランキング表示と発話履歴表示を MMDAgent へ追加した。

実験後のアンケート結果から、システムとの対話時にランキング・共有発話履歴を表示することは、話し掛けに対して高い効果があることがわかった。一方で、対話セット登録時にクリエイターや対話セットの統計情報を表示することは、高い効果がある項目もあったが低い項目もあった。低い項目については、おすすめの登録キーワードを提示することなどの、対話セット登録に対するインセンティブを高める方策が考えられる。また、提案システムは今後の期待度、総合評価が高かった。このことから、ユーザとクリエイターを相互に刺激することは、ユーザ生成型音声対話システムがユーザ生成型メディアとして発展させるための要素となることを示唆したと言える。

今後の課題としては、オープンな環境での大規模実験が挙げられる。

謝辞 本研究の一部は、科学技術振興機構 (JST) の戦略的基礎研究推進事業 (CREST) における研究領域「共生社会に向けた人間調和型情報技術の構築」の支援により行われた。また、本研究の一部は科研費 (21300066) の助成を受けたものである。

参考文献

- [1] 中川聖一:小特集-音声対話システムの実力と課題-小特集に寄せて-音声対話システム構築の課題-, 日本音響学会誌, 54 巻 11 号, pp.783-790(1998).
- [2] Wikipedia
<http://www.wikipedia.org/>
- [3] YouTube
<http://www.youtube.com/>
- [4] Oded Nov:What Motivates Wikipedians?, Communications of the ACM, Vol.50, NO.11, pp.60-64, (2007.11).
- [5] 鎌田麻以子, 加藤大志, 國枝和雄, 山田敬嗣: 創造連鎖活動における参加動機に関する調査, 情報処理学会ヒューマンコンピュータインタラクション研究会, Vol.2009-HCI-135, pp.1-8, (2009.11).
- [6] 鎌田麻以子, 加藤大志, 國枝和雄, 山田敬嗣: 共創活動の参加動機に関する調査-Wikipedia 参加者の参加動機因子-, 電子情報通信学会技術研究報告. SWIN, ソフトウェアインタプライズモデリング 109(430), pp.7-12, (2010.2).
- [7] 福田敏則, 吉見孔孝, 南角吉彦, 李晃伸, 徳田恵一: ユーザ生成型音声対話コンテンツを用いた音声情報案内システム, 電子情報通信学会技術研究報告. NLC, 言語理解とコミュニケーション 109 巻 355 号, pp.207-212, (2012.12).
- [8] 船谷内泰斗, 大浦圭一郎, 南角吉彦, 李晃伸, 徳田恵一: ユーザ生成型音声対話コンテンツに向けた有限状態トランスデューサに基づく簡潔な対話記述法の検討, 日本音響学会秋季研究発表会, 2-P-28, pp.223-224, (2013.9).
- [9] 三浦俊介: ユーザ生成型音声対話システムのための段階的なインセンティブ付与の検討, 卒業研究論文 (2011).
- [10] 李晃伸, 大浦圭一郎, 徳田恵一: 魅力ある音声インタラクションシステムを構築するためのオープンソースツールキット MMDAgent, 電子情報通信学会技術研究報告, 111 巻 364 号, pp.159-164(2011.12).
- [11] ニコニコ動画
<http://www.nicovideo.jp/>