

仮想エージェントを用いた オンラインサーベイ支援サービスの試作と予備的評価

高槻 大貴[†] 佐伯 幸郎[†] 中村 匡秀[†]

[†] 神戸大学 〒657-8501 神戸市灘区六甲台町 1-1

[†] 理化学研究所・革新知能統合研究センター 〒103-0027 東京都中央区日本橋 1-4-1

E-mail: [†]takatsuki@ws.cs.kobe-u.ac.jp, ^{††}sachio@carp.kobe-u.ac.jp, ^{†††}masa-n@cs.kobe-u.ac.jp

あらまし 本稿では、Formroid という仮想エージェント技術を用いてオンラインアンケート調査の回答が容易になる新しいシステムを紹介する。調査者によるアンケートの場合、Formroid は仮想エージェントに対して回答者に各質問に答えるように指示する。回答者は仮想エージェントとの会話を通じて、アンケートに回答することによって、Formroid は従来のフォーム入力を仮想エージェントによって行われる対面のインタビュー形式に変換する。本稿では、Formroid の設計の問題とプロトタイプの実装について述べる。また Formroid を用いたオンラインサーベイの実験と Quality Measurement Methods に基づいた予備的評価を行うことによって Formroid の満足度が確かめた。

キーワード ヴァーチャルエージェント, Web サービス

Prototyping and Preliminary Evaluation of Online Survey Support System Using Virtual Agent

Daiki TAKATSUKI[†], Sachio SAIKI[†], and Masahide NAKAMURA[†]

[†] Kobe University Rokkodai-cho 1-1, Nada-ku, Kobe, Hyogo, 657-8501 Japan

[†] Riken AIP, 1-4-1 Nihon-bashi, Chuo-ku, Tokyo 103-0027

E-mail: [†]takatsuki@ws.cs.kobe-u.ac.jp, ^{††}sachio@carp.kobe-u.ac.jp, ^{†††}masa-n@cs.kobe-u.ac.jp

Abstract In this paper, we present a novel system, called Formroid, which facilitates answering online questionnaire surveys with the virtual agent technology. For a questionnaire given by an investigator, Formroid commands the virtual agent to ask each question to the respondent. Through conversation with the virtual agent, a respondent can answer the questionnaire. Thus, Formroid transforms the conventional form input into a face-to-face interview conducted by the virtual agent. In this paper, we especially address the design issues of Formroid, and the implementation of prototype system. We also confirmed the satisfaction of Formroid by conducting online survey experiment using Formroid and preliminary evaluation based on Quality Measurement Methods.

Key words Virtual agent, Web service

1. はじめに

アンケート調査は、社会的な調査や研究において、基本的であるが重要な調査方法である。この方法は、例えば、国勢調査、市場調査、顧客満足度、製品の使用可能性テスト、健康管理評価、主観的実験などで広く使用されている。アンケート調査では、調査者は最初にアンケート用紙を作成する。この用紙は通常、調査に関連する複数の質問で構成されている。その後、フォームは回答者と呼ばれる目的の人々に配布される。各回答者はフォームの質問に答え、回答済みのフォームを調査者に返

す。伝統的に、アンケート調査は紙と鉛筆の道具を使って行われてきました。現在、インターネットの普及により、アンケート調査はオンラインで行われるようになりました。オンライン調査では、調査者は Web フォームと呼ばれる HTML 形式のアンケートフォームを作成する。Web フォームでは、質問はラジオボタン、チェックボックス、ドロップダウンリスト、テキストボックスなどの GUI コンポーネントによって構成されている。開発した Web フォームは Web サーバーにデプロイされ、フォームの URL が回答者に通知される。回答者は、PC または

スマートフォンを使用して Web ブラウザで URL を開き、各質問の値を選択（または入力）して、最後に [送信] ボタンをクリックして回答を送信する。フォームの値はサーバーに送信され、調査担当者によって指定されたデータベースに格納される。紙の調査と比較して、オンライン調査はアンケートの迅速な配達と収集を可能にし、手書きの回答を書き起こす手間を削減する。したがって、研究者は非常に効率的に研究と調査を行うことができる。有用なクラウドサービス（例えば、GoogleForms）もあり、それを使用することで、どのようなユーザーでも自分のオンラインアンケート調査用の Web フォーム簡単に開発することができる。調査者の利益と引き換えに、オンライン調査では、回答者は PC またはスマートフォンを使用して Web フォーム上で入力操作を行う必要がある。これは特に高齢者にとって精神的な障害となる。一般の回答者にとっても、長い Web フォーム上で入力操作は非常に単調で面倒である。Web フォームがテキストベースで表示されるため、調査へのユーザーの参加も少なくなる。問題は、回答者とのヒューマン - コンピュータインタラクションに対する Web フォームの能力の低さにある。もちろん、[4] [5] で研究されているように、調査研究へのユーザーの関与に影響を与えることができる他の多くの要因がある。しかし、本稿は特に、オンラインアンケート調査における人間とコンピュータの相互作用を強化することを目的としているので、より幅広い範囲の回答者が質問に答えることを楽しむことができる。そのために、Web フォームの代替技術として仮想エージェント [6] [7] を使用することを提案する。仮想エージェントは、PC 上で実行される、アニメーション化された、人間のようなグラフィカルチャットボットプログラムである。音声認識および合成技術を利用して、仮想エージェントは音声を通して人間のユーザーと通信することができる。私たちの重要なアイデアは、回答者とエージェントの間の会話で Web フォームを包むことである。本稿では、Formroid という新しいオンラインアンケート調査システムを紹介する。Formroid は各回答者の PC（またはスマートフォン）にインストールされることになっている。調査者は最初にアンケートフォームを作成し、そのフォームをクラウドデータベースに投稿する。Formroid は投稿されたアンケートデータを解釈し、仮想エージェントに各質問を回答者に尋ねるように指示する。与えられた質問ごとに、回答者は音声または単純なクリック操作で答えを入力する。入力に基づいて、仮想エージェントは回答者に回答し続けるように促すための反応を実行する。このように、回答者は仮想エージェントとの会話を通してアンケート全体を完成させる。Formroid は、収集した回答をクラウドデータベースに送信する。その結果、Formroid は従来のフォーム入力タスクを仮想エージェントによる対面インタビューに仮想的に変換する。本研究では、システムアーキテクチャ、データモデル、画面レイアウト、シーンなど、Formroid の設計上の問題に注目する。C # と Unity プラットフォーム [3] を使って、Formroid のプロトタイプ版も実装する。さらに、実装されたプロトタイプを用いて 2 種類のアンケートに回答してもらった実験を行い、Quality Measurement

Methods[12] に基づいてサービスの予備的評価を行う。

2. Formroid

2.1 System Architecture

図 1 に Formroid のシステムアーキテクチャを示す。この図では、調査者と回答者の 2 人のアクターがいる。調査者はアンケート調査を実施したい人であり、回答者はアンケートに回答する対象者である。クラウドには、フォームデータベースと回答データベースの 2 つのデータベースが配置されている。フォームデータベースは調査者によって作成されたアンケートフォームを格納し、回答データベースは回答者から送信された回答を格納する。2 つのデータベースには、外部アプリケーションが HTTP-POST メソッド（または HTTP-GET メソッド）でデータを挿入（または取得）できる Web-API がある。Formroid は、回答者の PC（またはスマートフォン）で実行されているアプリケーションソフトウェアである。Formroid はフォームデータベースから指定された質問表を取得する。その後、アンケートの構造を解析して一連の質問を識別する。各質問に対して、Formroid は仮想エージェントを介して回答者に尋ねる。回答者が音声または単純なクリック操作で質問に回答すると、Formroid は回答を記録する。すべての質問に回答すると、Formroid は記録された回答を回答データベースに送信する。

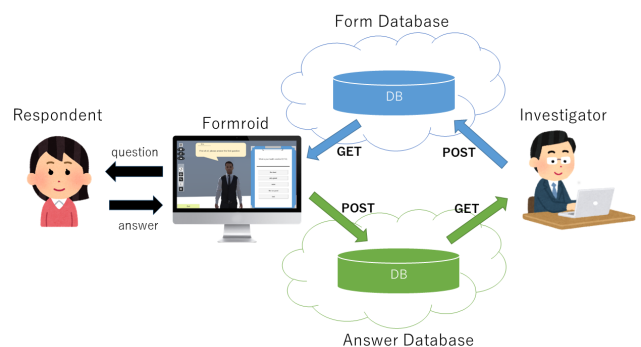


図 1 Formroid システムアーキテクチャ

2.2 データモデル

このサブセクションでは、アンケートフォームと回答データがデータベースによってどのように管理されるかについて説明する。まず、フォームのデータスキーマについて説明する。図 2 (a) は、JSON 形式で表されたフォームデータのインスタンスを示している。各フォームは 4 つの属性、id、name、creator、および questions から構成されている。id はフォームを識別する一意の文字列である。name と creator は、それぞれフォームの名前と作成者の識別子である。質問は一連の質問であり、アンケートの本文である。各質問は qid によって識別される。type は質問のタイプを定義し、以下の値のうちの 1 つによって指定される。・ 選択：与えられた選択肢から 1 つ選択する。・ チェック：選択肢内に当てはまる項目をすべてチェックする。・ 値：特定の値に答える。・ free：無料のコメントを説明

```
{
  "id": "LRAE01-SF12",
  "name": "SF-12 for life-rhythm assessment experiment 01",
  "creator": "longniu@ws.cs.kobe-u.ac.jp",
  "questions": [
    {
      "qid": "q1",
      "type": "select",
      "valueType": "int",
      "question": "What is your health condition?",
      "choices": [
        {"5": "the best"}, {"4": "very good"}, {"3": "soso"}, {"2": "Not so good"}, {"1": "bad"}
      ]
    },
    {
      "qid": "q2",
      "type": "select",
      "valueType": "int",
      "question": "Do you feel difficult to do light work for health reasons? For instance, clean your house and garden, or go out for a walk?",
      "choices": [
        {"3": "Very difficult"}, {"2": "A bit difficult"}, {"1": "Not difficult at all"}
      ]
    },
    ...
  ]
}
```

(a) フォーム

```
{
  "id": "20180710123456_5a3a5ab83",
  "timestamp": "2018-07-10T12:34:56",
  "respondent": "takatsuki@ws.cs.kobe-u.ac.jp",
  "form": "LRAE01-SF12",
  "answers": [
    {"qid": "q1", "answer": 3},
    {"qid": "q2", "answer": 4},
    {"qid": "q3", "answer": 4},
    {"qid": "q4", "answer": 1},
    {"qid": "q5", "answer": 2},
    {"qid": "q6", "answer": 3},
    ...
  ]
}
```

(b) 回答

図2 Data schema of Formroid (in JSON format)

する。valueType は、予想される回答値のタイプを定義する。question は質問の説明を定義する。choice は、質問に関連付けられた選択肢のセットである。選択肢は、回答値と説明のペアによって指定される。図2 (a) は、我々のプロジェクトにおける実験のために作成されたアンケート SF-12 [8] の一部を示している。最初の質問 q1 は回答者の健康状態を尋ねる。タイプが選択されているので、質問は回答者が「最良」から「悪い」までの5つの選択肢から1つを選択することを期待している。答えは整数値で返される。

次に回答のデータスキーマについて説明する。図2 (b) は、回答データの例を示している。各回答は、id: 回答の識別子、

timestamp: 回答が収集された時刻、回答者: 回答者の識別子、フォーム: 対応するフォームの識別子、および回答: 収集された回答の配列で構成される。各回答は、質問 ID と回答値のペアによって指定される。図2 (b) は、それが 2018-07-10T12:34:56 に takatsuki @ ... によって応答された LRAE01-SF12 形式の答えであることを表している。質問 q1, q2, ..., q6 の回答値は、それぞれ 3, 4, 4, 1, 2, 3 である。

2.3 画面構成

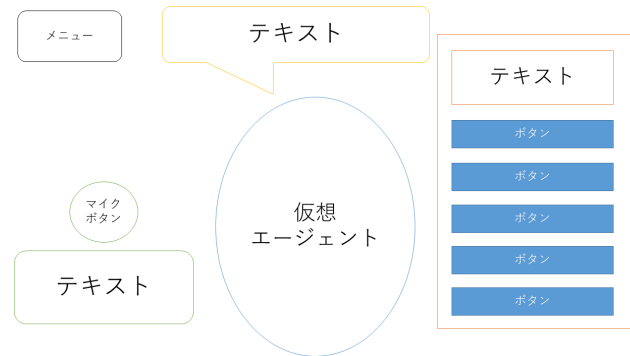


図3 Formroid 画面レイアウト

図3に示す Formroid の画面レイアウトを設計する。仮想エージェントが画面の中央に配置され、各質問を回答者に話す。スピーチはまた上記の吹き出しのテキストで示されている。質問の選択肢は右側に表示され、各選択肢はボタンとして表される。質問に答えるために、回答者は左側のマイクボタンをクリックし、3秒間の音声認識する間に選択肢または選択番号を話す、またはタッチスクリーンディスプレイが利用可能なときにボタンに触れることで回答できる。また音声認識された結果を左したにテキストで表示される。画面の左上には、回答者がアンケートや Formroid を設定できるメニューツールが列挙されている。設定には、仮想エージェントの音声の有無や Formroid のモード選択などが含まれる。

2.4 アンケート調査のシーン遷移

Formroid の動作はシーンによって構成されている。シーンは、仮想エージェントが回答者とどのような順序で対話するかを定義する Formroid のビルディングブロックである。Formroid の現在のバージョンは以下のシーンで構築されている。

(Scene 1) はじめに: 仮想エージェントは、応対者に挨拶し、簡単な指示をする。

(Scene 2) 名前と日付の確認: 仮想エージェントは回答者の名前と調査日を確認する。正しくない場合、回答者は正しい名前と日付を伝える。

(Scene 3) 調査の選択: フォームデータベースから、回答者に適用される調査のリストを取得する。次に、仮想エージェントは、回答者がどの調査に回答したいかを尋ねる。回答者が1つを選択すると、仮想エージェントは回答者が開始の準備がきているかどうかを確認する。

(Scene 4) 調査開始: Formroid は選択されたフォームの JSON データをフォームデータベースから取得し、それから一

連の質問を抽出する。仮想エージェントは調査セッションを開始するように指示する。

(Scene 5) 質問をする：各質問について、仮想エージェントは質問を話し、画面に選択肢を表示する。回答者は、該当する選択項目を話すか、対応するボタンをクリック（またはタッチ）して質問に答える。Formroid は質問に対する答えを記録する。インタラクション中、仮想エージェントは、調査の進行状況と前の回答に基づいて、動画の動きに対する反応を示す。反応は、例えば、「素晴らしい！ 10 番目の質問に達しました。あなたは調査の半分をしました！ 立ち止まるな！」。これらの反応は、回答者の関与を高めるコミュニケーションを促進することを目的としている。

(Scene 6) 調査終了：全ての質問が終了すると、仮想エージェントは調査終了を指示する。記録されたすべての回答から、Formroid は回答データを JSON 形式で作成し、JSON を回答データベースに投稿する。

(Scene 7) エピローグ：仮想エージェントは、回答者の調査に対する努力を評価し、調査セッションを終了する。

各回答者用にパーソナライズされた調査セッションを達成するために、Formroid は動的に上記のシーンを作成する。より具体的には、上記の各シーンについて、Formroid はシーンテンプレートを管理する。シーンの一部はオープン変数である。回答者の氏名、日付、および指定されたアンケート調査に基づいて、Formroid は変数に具体的な値を割り当てる。これにより、実行時にパーソナライズされたシーンが構築される。

3. プロトタイプ開発

Formroid のプロトタイプ版を実装しました。プロトタイプに使用されたテクノロジーは次のとおりである。

- 言語: C# 7.2
- シーンエンジン: Unity Personal 2018.2.8f1
- エージェント 3D モデル: ユニティちゃん 3D モデルデータ

Unity は 3D / 2D オブジェクトやサウンドを操作するための強力なツールキットを持っているので、私たちは Formroid のシーンを実装するために Unity [3] を使いました。また、すぐに使えるデータリソースであるさまざまな資産もある。プロトタイプでは、仮想エージェントのアバターのために 3D ヒューマンモデルの無料資産を借りている。Unity が 3D モデルを他のモデルに置き換えるのは簡単である。つまり、エージェントのアバターを簡単に変更できる。

4. 予備的実験

4.1 実験概要

本章では、試作した Fomroid の満足度や価値を評価するために、予備的評価実験を行う。評価実験には、名前や生年月日、性別などの入力系のアンケートと 2 択の中から 1 つ選ぶ選択系のアンケートの 2 種類のアンケートを用いて実験する。本実験では、5 名の被験者に Formroid のプロトタイプを用いて、2 種



図 4 Formroid プロトタイプ

類のアンケートに回答してもらった後、Fomroid を使用してみてもう一度か等についてのアンケートを行い、サービスを評価した。

4.2 評価指標

本実験では、開発したサービスがユーザの満足度をどれだけ満たしていたかを調べるために、Quality Measurement Methods[12] を利用して評価を行う。Quality Measurement Methods とは、早稲田大学グローバルソフトウェアエンジニアリング研究所が 2015-2016 年度に RISE 委託研究 2015-16 年度「測定評価と分析によるソフトウェア製品品質の実態定量化および総合的品質評価枠組みの確立」を実施した成果を、国際規格 ISO/IEC 25000 SQuaRE シリーズに基づくソフトウェア製品の品質測定評価枠組み、および、同枠組みを用いて日本の 21 製品に対する測定評価結果をまとめたものである。Quality Measurement Method でサービスの品質を評価する際に、表 1 の品質モデルを用いる。

今回は利用時の品質について満足性の評価する。以下に評価を行う満足性の特性の詳細について述べる。

満足性：満足性は、システム利用によって利用者のニーズが満足される度合いと定義されている。システム利用によって利用者がどの程度満足するかを評価する。副特性の有用性・信用性・快感性・快適性については、それぞれ次の点について評価を行う。有用性は、利用者が行いたいことが達成され、満足するかを評価する。信用性は、利用者がシステムの動作として提示され、期待して、行いたいことが達成されるという確信をどの程度持てるかを評価する。快感性は、アンケートをする中で楽しんで回答できたかを評価する。快適性は、システム操作上の快適性を提供できるかを評価する。

Fomroid の満足度に対する質問項目と各質問に対する回答の選択肢を次に示す。各質問は 1～5 の五段階で評価される。選択肢 1 が最低値の評価 1 となり、選択肢 5 が最高値の評価 5 となる。

Satisfaction

(Sub)characteristics	SQuaRE	ID	Name	Definition	Details
Usefulness	SUs-1	SUs.1.1	Degree of satisfaction for the product	$X' = \text{Mean value of the user response in } X$	$X = \text{Response to a user questionnaire related to user satisfaction}$
Usefulness	SUs-1	SUs.1.2	Net promoter score	$X' = \text{Mean value of the user response in } X$	$X = \text{Response to a user questionnaire related to the net promoter score}$
Usefulness	Sus-2	Sus.2.1	Degree of satisfaction for the functionalities	$X' = \text{Mean value of the user response in } X$	$X = \text{Response of the mean satisfaction by a user in the user questionnaire related to satisfaction of each functionality}$
Trust	STr-1	STr.1.1	Degree of trust	$X' = \text{Mean value of the user response in } X$	$X = \text{Response to a user questionnaire related to trust}$
Pleasure	SPI-1	SPI.1.1	Degree of pleasure	$X' = \text{Mean value of the user response in } X$	$X = \text{Response to a user questionnaire related to stress-free use}$
Pleasure	SCo-1	SCo.1.1	Degree of pleasure	$X' = \text{Mean value of the user response in } X$	$X = \text{Response to a user questionnaire related to pleasure}$

表 1 Quality Measurement Methods

質問項目

1. 好感はもてましたか
2. 役に立つと思うか
3. アンケートの回答に使いたいと思うか
4. Net 技術がアンケートを支援しているとおもいるか
5. アンケート機能は満足いくものだったか
6. 音声認識や音声合成は満足いくものだったか
7. 信頼して使うことができましたか
8. 操作はわかりやすいであるか
9. 大抵の人が使えると思うか
10. 画面は見やすいであるか
11. 楽しくアンケートに回答することができましたか

選択肢

1. まったく思わない
2. あまり思わない
3. 普通
4. 少し思う
5. とても思う

質問 1・2・3・4・5・6 までの質問で有用性を評価し、質問 7 で信頼性、質問 8・9・10 で快適性、質問 11 で快感性を評価する。

4.3 実験結果

本実験では 5 名の被験者に Formroid のプロトタイプを用いて 2 種類のアンケートに回答してもらった後、Formroid の満足性についての評価アンケートを 5 段階で回答してもらった。以下に各質問ごとの評価結果と副特性である有用性・信頼性・快適性・快感性ごとの評価結果を図 5 と図 6 に示した。

満足性の質問ごとの評価

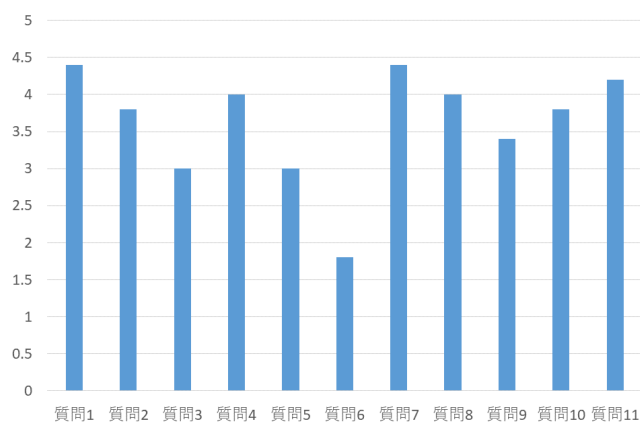


図 5 満足性の各質問ごとの評価結果

図 5 は横軸が各質問で縦軸は被験者 5 人の 5 段階評価の平均

満足性の副特性ごとの評価

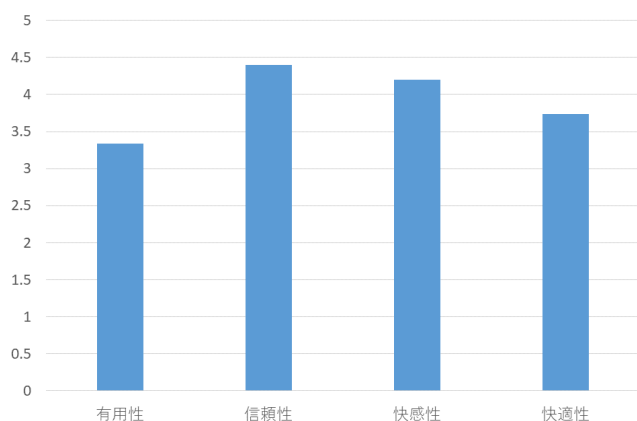


図 6 副特性ごとの評価結果

値である。質問ごとの評価では質問 1「好感は持てましたか」と質問 7「信頼して使うことができましたか」が最も平均評価が高く平均値は 4.4 となった、反対に質問 6「音声認識や音声合成は満足いくものだったか」は平均値が 1.8 となり最も低い結果となった。質問 6 以外の評価では平均値が 3 以上の評価を得ており、質問 6 のみが 3 よりも低い評価となった。本実験の最後に Formroid を使ってみての感想・意見を記入してもらったところ、以下の点が挙げられた。

感想・意見

- 時間はかかるがそれで苦痛になることはなかった
- 口にだしてしゃべることで気が紛れている気がする
- 音声入力 of 誤認識が多い
- 音声を入力してから認識するまでの時間が少しくになる
- スピーカーの音声合成が音声認識に反応してしまう
- 名前や生年月日など記入するものを音声認識で回答することは難しいが、選択をする方は楽に回答できた

特に音声認識の精度が良くないといった感想を多くの被験者が挙げていた。これらの感想・意見から、音声認識を用いて回答することへの負担は少ないが、音声認識の誤認識による負担が特に大きいため質問 6「音声認識や音声合成は満足いくものだったか」の評価が低くなったと考えられる。

図 6 は横軸が副特性で縦軸は副特性ごとの各質問の平均値を示したものである。副特性ごとの評価ではどの副特性でも普通以上の平均的評価を得ていることがわかる。特に信頼性と快感性の平均値は 4 を超えている。また有用性が一番低い結果となっているが、音声認識の課題を解決することで Formroid の有用性を向上させることができると考えられる。

5. おわりに

私たちは今回 Formroid と呼ばれる仮想エージェントを用いた新しいオンラインサーベイ支援システムの試作と予備的評価を紹介した。本稿では Formroid の設計問題とプロトタイプを

扱った。また Formroid のプロトタイプを用いて 5 名の被験者に 2 種類のアンケートを回答してもらい、Quality Measurement Methods に基づいた満足性の予備的評価をした。現在、私たちは他の可能なユースケースとフィードバックを考慮して、実装の改良をしている。特に音声認識の機能は改良する予定である。今後の課題といたしまして、Formroid の利点と欠点を正しく評価するために大多数のさまざまな年齢や性別の人から評価をする必要がある。また評価指標として満足性だけでなく、有効性や効率性などの観点からの評価も考慮した評価実験を検討する。

謝辞 この研究の一部は、科学技術研究費（基盤研究 B 16H02908, 15H02701, 基盤研究 A 17H00731, 萌芽研究 15K12020）、および、立石科学技術振興財団の研究助成を受けて行われている。

文 献

- [1] L.A. Suskie, Questionnaire Survey Research: What Works, Resources for Institutional Research, Association for Institutional Research, 1996.
- [2] “Google forms”. (Date last accessed 13-September-2018). <https://docs.google.com/forms>
- [3] “Unity”. (Date last accessed 13-September-2018). <https://unity3d.com>
- [4] A. Saleh and K. Bista, “Examining factors impacting online survey response rates in educational research: Perceptions of graduate students,” Journal of MultiDisciplinary Evaluation, vol.13, no.29, pp.63–74, 2017.
- [5] D.D. Nulty, “The adequacy of response rates to online and paper surveys: what can be done?,” Assessment & Evaluation in Higher Education, vol.33, no.3, pp.301–314, 2008.
- [6] S. Tokunaga, K. Tamamizu, S. Saiki, M. Nakamura, and K. Yasuda, “VirtualCareGiver: Personalized smart elderly care,” International Journal of Software Innovation (IJSI), vol.5, no.1, pp.30–43, Oct. 2016. DOI: 10.4018/IJSI.2017010103, <http://www.igi-global.com/journals/abstract-announcement/158780>.
- [7] H. Horiuchi, S. Saiki, S. Matsumoto, and M. Nakamura, “Virtual agent as a user interface for home network system,” International Journal of Software Innovation, vol.3, no.2, pp.24–34, April 2015.
- [8] J.E. Ware, M. Kosinski, and S.D. Keller, “A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity,” Medical Care, vol.34, no.3, pp.220–233, 1996. <http://www.jstor.org/stable/3766749>
- [9] J.E. Ware, M. Kosinski, J.E. Dewey, and B. Gandek, How to score and interpret single-item health status measures: a manual for users of the SF-8 Health Survey, Quality-Metric Incorporate, 2001.
- [10] L. Niu, S. Saiki, and M. Nakamura, “Using non-intrusive environmental sensing for ADL recognition in one-person household,” International Journal of Software Innovation (IJSI), vol.6, no.4, pp.16–29, Aug. 2018.
- [11] L. Niu, S. Saiki, and M. Nakamura, “Integrating environmental sensing and BLE-based location for improving ADL recognition,” The 19th International Conference on Information Integration and Web-based Applications & Services (iiWAS2017), pp. ••–••, Dec. 2017. Salzburg, Austria.
- [12] “Ipa rise 委託研究 2015-16 年度測定評価と分析によるソフトウェア製品品質の実態定量化および総合的品質評価枠組みの確立”. (Date last accessed 14-February-2019). <http://www.washi.cs.waseda.ac.jp/wsqb/>