

用例評価のモチベーション維持支援システム 「用例の森」の開発と評価

狩野 翔^{1,a)} 福島 拓^{1,b)} 吉野 孝^{2,c)}

受付日 2011年4月13日, 採録日 2011年10月3日

概要: 現在, モチベーション維持に関する研究が行われている. しかし, 「日常的に行わない」作業を対象とした支援は十分に考慮されていない. 「日常的に行わない」作業では, 「日常的に行う」作業に対して存在する「やらなければならない」という気持ちが支援対象者の中に少なく, モチベーション維持を支援することは難しい. そこで本論文では, 非日常的な作業として「医療分野向け用例の評価」を対象とし, そのモチベーション維持支援を行った. 医療分野向けの用例では, 医療現場へ提供するために正確性の評価が必要である. しかし, 用例数が多く作業のゴールが見えない, 評価作業が単調で飽きるといった問題が存在する. そのため, 「楽しさ」や「達成感」の要素を用いた用例評価のモチベーション維持支援システム「用例の森」を開発し, 評価を行った. 本論文の貢献は次の4点にまとめられる. (1) 非日常的な作業に対するモチベーション維持支援システム「用例の森」の提案を行い, 実現した. 「用例の森」は, 非日常的な作業である用例の評価活動において, モチベーション維持支援の可能性があることを示した. (2) 用例を木に, 用例の評価状況を木の成長に見立てることにより, ユーザに「楽しさ」や「達成感」を与えることができ, 「評価状況の把握」に役立つことが分かった. また, 用例群をいくつかのグループに分け, 小さな目標としてユーザに提示することにより, 達成感を与えることの補助ができることが分かった. (3) ユーザは自分の登録した用例が評価されると, 他のユーザが登録した用例を評価したいと思う傾向があることを明らかにした. (4) ユーザは手軽に達成感を得るために, 数の少ない用例群を優先して評価する傾向があることを明らかにした.

キーワード: コミュニティ支援, 共有仮想環境, 協調作業支援, モチベーション支援, 用例評価

Development and Evaluation of the Sustainable Motivation Support System “Example-sentence Forest” for the Evaluation of Example Sentences

SHOU KARINO^{1,a)} TAKU FUKUSHIMA^{1,b)} TAKASHI YOSHINO^{2,c)}

Received: April 13, 2011, Accepted: October 3, 2011

Abstract: Recently, scholars have been studied supporting sustainable motivation. However, these studies do not support for “non-daily works.” In the context of “non-daily works”, it is difficult to apply the sustainable motivation support system because people do not have think that “must do.” In our study, we support users’ motivation for “Evaluating example sentences in the medical field.” In the medical field, example sentences need to evaluate the accuracy of sentences. However, problems related to the evaluation of example sentences have existed; for instance, there is no anticipated end to the evaluation work because there are too many sentences, the work is tedious, and so on. Therefore, we have developed and evaluated a sustainable motivation support system, “Example-sentence Forest” that is used for the evaluation of example sentences employs factors like “fun” and “accomplishment.” The contributions of this paper are the following results: (1) We proposed a sustainable motivation support system, this is, “Example-sentence Forest” for non-daily works. The system has a possibility of fostering motivation in users for evaluating works that are non-daily works. (2) It is useful for experiencing a sense of fun and accomplishment and assessing the evaluation that use example sentences resemble to “Trees”, and use evaluated status resemble to growth of the Tree. We divided the example-sentences groups into some to present them to users as a small target. We found that a small target can give the sense of accomplishment. (3) We revealed that the user tend to evaluate sentences that drew up by other user when my sentences are evaluated. (4) We revealed that the user tend to evaluate sentence groups that have few sentences.

Keywords: community support, shared virtual environment, collaboration support, motivation support, evaluating example-sentences

¹ 和歌山大学大学院システム工学研究科
Graduate School of Systems Engineering, Wakayama University, Wakayama 640-8510, Japan

² 和歌山大学システム工学部
Faculty of Systems Engineering, Wakayama University,

Wakayama 640-8510, Japan

a) karino@yoslab.net

b) fukushima@yoslab.net

c) yoshino@sys.wakayama-u.ac.jp

1. はじめに

現在、日常の作業や生活習慣、学習などを対象としたモチベーション維持支援に関する研究が数多く行われている [1], [2], [3], [4]. これらの研究では、情報技術を用いて日常の作業や学習を意欲的に行わせることなどを目的としている [3], [4]. しかし、これらの研究では用例の評価活動などの「日常的に行わない」作業を対象としたモチベーション維持は十分に考慮されていない. 支援対象者にとって日常的な作業や学習は、必然的な行動、または自分に直接利益のある行動であり、「やらなければならない」という気持ちが支援対象者の中に存在する. この気持ちを刺激、または新しい要素を付加することで、これらのモチベーション維持の支援が可能となっている. しかし、「日常的に行わない」作業では、「やらなければならない」という気持ちが支援対象者の中にあまり存在せず、自分に直接利益のない場合が多い. そのため、作業に興味がある支援対象者以外のモチベーション維持を支援することは難しい. 「日常的に行わない」作業では、「日常的に行う」作業の利益である報酬や学習結果、自身の健康といった要素以外を用いてモチベーション維持の支援を行う必要がある.

本論文では、エンタテインメント要素を用いることにより、「日常的に行わない」作業に対するユーザのモチベーション維持支援を目指す. また、支援の対象として「医療分野向け用例の評価活動」に着目した. 現在、外国人登録者数は増加傾向にあり [5], 日本語を理解できない外国人が多数存在している [6]. そのため、特に医療分野では言葉の違いによる医療ミスが発生する恐れがある. そこで、医療分野への提供を目的とした、多言語用例対訳の収集、共有に関する研究が行われている [7], [8]. 用例対訳とは、あらかじめ翻訳された同じ意味の用例である. 収集された用例対訳を医療現場に提供するためには、各用例の正確性を評価する必要がある. しかし、評価する必要がある用例の数は多く、評価作業のゴールが見えないため、単調で飽きるといった問題が存在している. また、評価を行うユーザは「ボランティア」であるため、用例の評価に対して興味を持っていないユーザのモチベーションは低い.

そこで本論文では、「楽しさ」や「達成感」などの要素を用いた、用例評価のモチベーション維持支援システム「用例の森」を提案する. 本システムは、用例1つ1つを木に見立てた仮想空間上の森を育成するシステムである. ユーザが用例を評価することで、用例に関連付けられた「用例の木」が成長していく. 木の成長を見ることで得られる楽しさや達成感により、用例の評価活動に対するユーザのモチベーションを維持することを目的とする.

本論文では、2章で関連研究について述べる. 3章では本システムの構成と機能について、4章では従来システム

との比較実験の結果と考察について述べる. 5章ではモチベーション維持効果検証のための実験の結果と考察について、最後に6章において本研究の結論を述べる.

2. 関連研究

本章では、これまでに行われてきた、作業に対するモチベーション維持に関する研究について述べる. 本研究の位置付けを明らかにするために、我々が考えているモチベーション要素と支援システムの関係モデルを示し、モデルをもとに既存研究と本研究の関係を明らかにする. 図1にモチベーション要素と支援システムの関係モデルを示す. また、表1に既存の支援システムと本システムをモデルに適用した例を示す.

支援対象者は作業に対して元々、いくつかのモチベーション要素を持っている (図1(1)). このモチベーション要素は作業によって異なる. 支援システムは支援対象者が持つモチベーション要素を強化したり、新たなモチベーション要素を加えたりすることで (図1(2)), 作業に対するモチベーション維持の支援を行う (図1(3)). また、システムのユーザインタフェースは、支援対象者が持つモチベーション要素をどれだけ作業に反映できるかの役割を果たす. たとえば、「使いにくい」ユーザインタフェースは、支援対象者が持つモチベーションを阻害する可能性がある. 作業に適したユーザインタフェースを支援対象者に提供することで、支援対象者が持つモチベーション要素を作業に対し、適切に反映することができる.

これらをもとに、まず、エンタテインメント要素を用いて支援を行う研究について述べる. エンタテインメントなどによる心理的な効果は、労働を行う人のモチベーション向上に役立つことが知られている [9]. たとえば、日常的な作業にエンタテインメント性を持たせて、モチベーション維持向上を行う研究として倉本らの研究がある [3]. 倉本らの開発したシステムである懐優館 [10] は、ユーザの主観的作業量に応じて餌を入手し、熱帯魚を成長させるものである. 他のユーザの熱帯魚と主観的に比較し、「あの人に勝った」「あの人に勝ちたい」とユーザそれぞれに思わせること

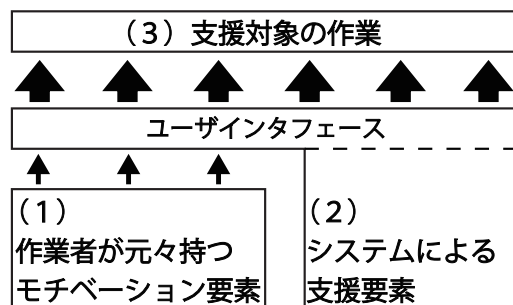


図1 モチベーション要素と支援システムの関係モデル
Fig. 1 Relational model between motivation elements and support system.

表 1 図 1 のモデルにおける既存システムと本システムの比較

Table 1 Comparison based on Fig.1 among the existing systems and our system.

| | 懐優館 [10] | Wikipedia | 本システム |
|------------------------------|--|--|-----------------------|
| (1) 作業者が元々持つ モチベーション要素 | ・やらなければならない 気持ち ・報酬 ・名声 ・競争 ・技術の熟達意欲 など | ・利他的行為 ・相互作用 ・連帯感 ・他者からの評価 ・コンテンツ作成 ・コンテンツ整理 など | ・利他的行為 ・コンテンツ評価 など |
| (2) システムによる 支援要素 | ・楽しさ ・達成感 ・他者との比較 など | ・使いやすさ ・素早いナビゲーション など | ・楽しさ ・達成感 など |
| (3) 支援対象の作業 | 日常のデスクワーク [10] (日常的な作業) | 記事作成 [12] (非日常的な作業) | 用例の評価 (非日常的な作業) |

・表中の (1), (2), (3) は図 1 の (1), (2), (3) にそれぞれ対応する。
 ・(1) の項目の一部は文献 [11], [12], [13] を参考にしている。

で、モチベーション維持を支援している。文献 [10] では、懐優館を日常的な作業であるデスクワークに適用し、モチベーション維持効果の検証を行っている。デスクワークを支援する場合、支援対象者が元々持つモチベーション要素として、日常的な作業の特徴である「やらなければならない気持ち」や「報酬」、また「成果による名声」「他者との競争」「技術の熟達意欲」など、多くの要素があてはまる (表 1(1))。また懐優館では、熱帯魚の育成による「楽しさ」「達成感」「他者との主観的な比較」により、モチベーション維持を支援している (表 1(2))。

次に、オンラインコミュニティの利用者のモチベーション分析に関する研究について述べる。山田らは、オンラインコミュニティにおけるモチベーション維持方法のモデル化を行っている [11]。また、Kuznetsov はオンライン百科事典 Wikipedia *1 における、記事作成者のモチベーションについて分析を行っている [12]。Wikipedia における記事作成作業は、「日常的に行わない」非日常的な作業といえる。Kuznetsov は記事投稿者のモチベーションとして「利他的行為」「他の利用者との相互作用」「コミュニティによる連帯感」「他のユーザから評価」などをあげている (表 1(1))。また、山田らの分類している 7 つのモチベーションのうち、「A: コンテンツを作りたい、見せたい欲求」「B: コンテンツを集めてまとめたい、整理したい欲求」などの要素もあてはまる。また、Kuznetsov は Wikipedia の使いやすさと素早いナビゲーションにより、これらのモチベーション要素を刺激していると述べている (表 1(2))。

最後に、本論文で支援する用例の評価活動に関するモチベーションについて、図 1 のモデルを利用して説明する。用例評価者が元々持つモチベーションとして、「利他的行為」や、山田らの分類している 7 つのモチベーションのう

ち「D: コンテンツを評価したい、コメントしたい欲求」などの要素があてはまる (表 1(1))。また、用例の評価活動は非日常的な作業であり、図 1(1) の「作業者が元々持つモチベーション要素」が、懐優館などのシステムで支援する日常的な作業よりも少ない。加えて、用例の評価活動は単調な作業であり、Wikipedia の記事作成者が持っている「コンテンツを作りたい、見せたい欲求」などのモチベーション要素を、用例評価者はあまり持っていない。さらに、Pink は「単純な作業に対しては報酬を用いることでモチベーションの維持が行える」と述べているが [13]、評価を行うユーザは「ボランティア」であるため、報酬などのモチベーション要素をユーザが持っていない可能性がある。これらのため、既存のモチベーション支援の仕組みでは、用例の評価に対するモチベーションを十分に維持できるかどうかは不明である。そこで本論文では、既存の研究で用いられてきた「楽しさ」や「達成感」などの要素による支援 (表 1(2)) を、図 1(2) の「システムによる支援要素」とする。非日常的な作業であり、支援対象者が元々持っているモチベーション要素が少ない「用例の評価活動」にこれらを適用することで、モチベーション維持効果を十分に得られるかどうか、検証を行う。

3. 用例の森

3.1 設計方針

本システムは用例評価活動のモチベーション維持を目的としたシステムである。用例の評価には、(1) 複数人による評価が必要、(2) 評価作業は単調であり、ゴールが見えない、という問題点が存在する。そこで本システムでは、(1) 木を成長させるという目的をユーザに与える、(2) 木の成長による楽しさや達成感を利用する、ことで問題点の解決を狙う。

本システムの設計方針を以下に示す。

*1 日本語メインページ: <http://ja.wikipedia.org/wiki/> (2011 年 10 月 8 日確認)

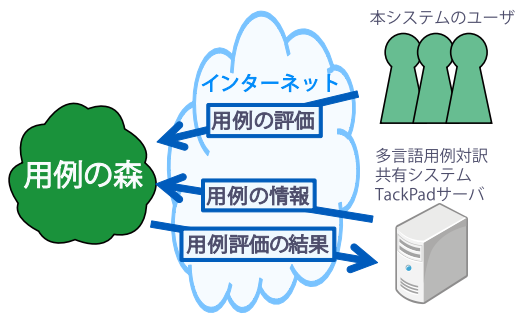


図 2 システム構成

Fig. 2 System configuration.

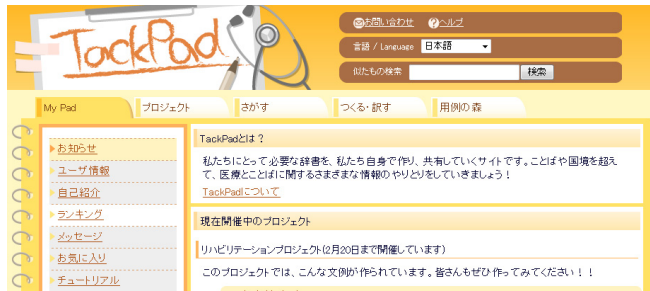


図 3 TackPad の画面例

Fig. 3 Screenshot of TackPad.

(1) 仮想空間を用いた楽しさや達成感の実現

用例を木に見立て、仮想空間上に森として表示する。木の成長をユーザに見せることで楽しさや達成感を与える。また、用例の評価状況を木の成長に見立てることで、評価状況の把握に役立たせる。森の育成による楽しさと、評価状況の把握による目標までの現状把握により、ユーザの継続的な評価活動を支援する。

(2) TackPad [7]*2との連動

多言語用例対訳共有システム TackPad と連動させる。TackPad において収集された用例に対する評価を支援する。

3.2 システムの構成

本システムの構成を図 2 に示す。本システムは多言語用例対訳共有システム TackPad の一機能 (Web ページ) として動作する。本システムは TackPad のデータベースから用例の情報を取得し、画面上に表示する。ユーザによる用例評価の結果は TackPad のデータベースに送信し、保存される。なお、開発は Flash (Adobe Systems) を利用して開発を行った。

3.3 多言語用例対訳共有システム TackPad

多言語用例対訳共有システム TackPad の画面例を図 3 に示す。TackPad は我々が開発した、医療分野の用例対訳を共有するシステムである [7]。TackPad におけるユーザ

は医療従事者や診察を受けた患者、翻訳者であり、アカウントを登録することでシステムを利用できる。登録したユーザは、医療分野に關係する用例 (TackPad では文例と呼ばれる) の登録や閲覧、評価を行うことができる。TackPad において用例の評価は「おすすめポイント」と呼ばれ、「病院でよく使う・使わない」「書き言葉・話し言葉」の 2 軸の評価で行う。

3.4 利用の流れ

本システムは、「森選択画面」「場所選択画面」「用例評価画面」の 3 つの画面から成り立っている。ユーザは以下の流れで本システムを利用する。

(1) 訪問したい森の選択 (森選択画面)

本システムにアクセスすると、初めに「森選択画面」が提示される。本システムでは、「用例に付与されているタグ」、または「用例を登録したユーザ」*3によって用例を分類し、1 つの「森」として扱う。たとえば図 4 のように、TackPad で「ノイローゼ」というタグがついている用例群は、本システムで「ノイローゼの森」と表記される。ユーザは本画面で「訪問したい森 = 評価したい用例のタグ」を選択する。

(2) 場所の選択 (場所選択画面)

(1) において訪問したい「森」を選択すると、「場所選択画面」が提示される (図 5)。「森」には用例が多数あり、1 つの画面にすべての用例を表示することは難しいため、森をいくつかの「場所」に分けることにしている。場所は図 5(A) のように、マスとして表示している。ここでユーザは「場所」を選択する。また、マスには図 5(B) の「その場所の中での現ユーザの評価率」を表すアイコンアニメーションが表示される。たとえば、1 つの場所に含まれる用例の評価率が 0% であれば種々のアイコン、100% であれば、花のアイコンが表示される。このような表示アイコンの変化によって、ユーザに達成感を与えることを狙っている。

(3) 用例の評価 (用例評価画面)

(2) の「場所」を選択すると、「用例評価画面」が提示される (図 6)。ここで、ユーザは用例の評価を行う。本システムでは、用例 1 つ 1 つを「用例の木」として扱っている。利用中のユーザが評価していない用例の木は半透明で表示される (図 6(A))。ユーザが用例の木をクリックすると、評価軸が現れる (図 6(B))。評価軸は TackPad で用いられているものと同じである。ユーザが木に関連付けられた用例の評価を行うことで、「用例の木」は木として成長する (図 6(C))。1 人のユーザによる同じ用例に対する評価は 1 回のみであるため、多数のユーザが用例を評価しなければ、「用

*2 <http://med.tackpad.net/>

*3 この機能は 5 章の実験で追加した機能である。

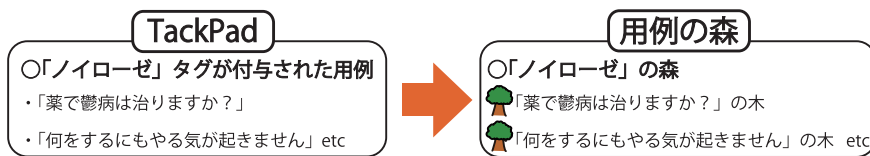


図 4 森の概念

Fig. 4 Concept of "Forest."

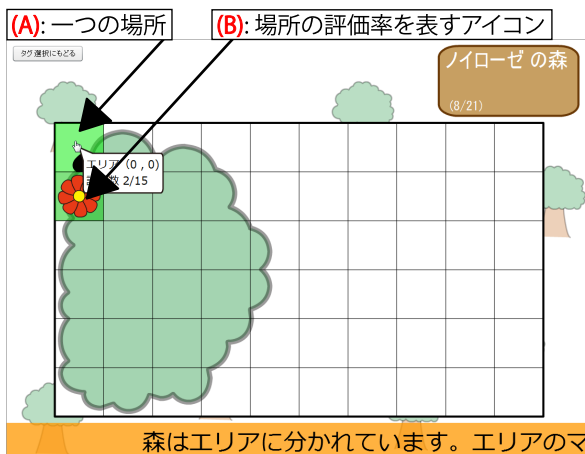


図 5 場所選択画面の例

Fig. 5 Screenshot of "Select place scene."

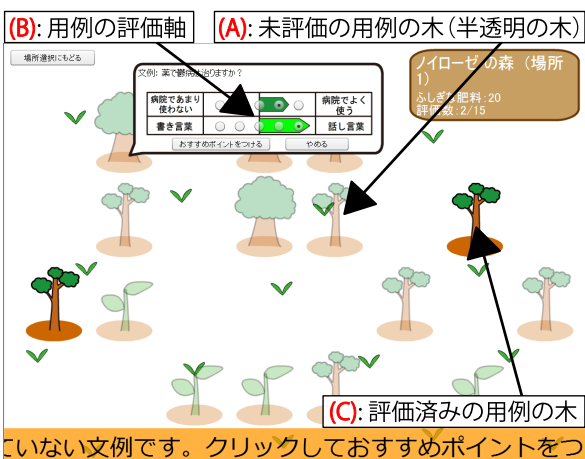


図 6 用例評価画面の例

Fig. 6 Screenshot of "Example sentence evaluation scene."

例の木」は大木に成長しない。花が咲いた木は「十分に評価された」用例となる。本システムでは、花が咲いた木になるまでに必要な評価数を5とした。これは TackPad において「5名の評価者に評価された用例は医療現場に提供可能」としているためである*4。

3.5 用例の森の機能

本システムにおいて利用できる機能について説明する。なお、これらの機能は5章の実験から利用できるようにした。

*4 4章の比較実験時には、暫定的に4名の評価により木の成長限界に達するようにしていた。

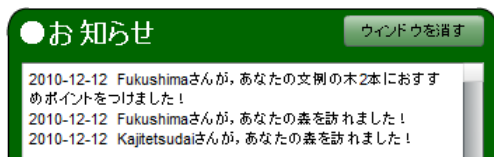


図 7 お知らせの例

Fig. 7 Screenshot of "Notice."

3.5.1 あなたの森

利用中のユーザ自身が登録した用例群である「あなたの森」という森を用意した。あなたの森では3.5.2項で説明する「お知らせ」を閲覧できるほか、自分が登録した用例の木がどれだけ成長したか確認することができる。自分が登録した用例の木が成長している様子を観覧することで、森の成長を楽しんでもらうこと、また他のユーザが登録した用例の木の評価に興味を持ってもらうことが狙いである。

3.5.2 お知らせ

お知らせの例を図7に示す。お知らせはあなたの森で閲覧できる。お知らせには「あなたの森(利用中のユーザの森)を訪れたユーザ」および「自分が登録した用例を評価したユーザ」が表示される。これは「訪問や、評価してくれたユーザの用例を評価しよう」と思ってもらうことを目的とした機能である。

4. 従来システムとの比較実験

本システムを用いた用例の評価活動と、TackPadの評価機能(以下「従来システム」と表記する)を用いた用例の評価活動とを比較する実験を行った。本実験の目的は、ユーザが用例の評価を行うにあたって、本システムと従来システムのどちらを好むかを検証することである。また、本システムを利用することで「楽しさ」や「達成感」を得られるかどうか、「評価状況の把握」ができるかどうかについての検証も行う。

4.1 検証項目

本実験では、以下の3つの項目について検証を行う。

- (1) ユーザが用例の評価を行うにあたって、本システムと従来システムのどちらを好むか。
- (2) 本システムを利用することで「楽しさ」や「達成感」を得られるか。
- (3) 本システムを利用することで「評価状況の把握」ができるか。

表 2 比較実験での用例のタグ

Table 2 Tags of example sentences in the comparison experiment.

| タグ名 | 用例の数 |
|-------|------|
| ノイローゼ | 7 |
| 切り傷 | 4 |
| 受付 | 28 |
| 歯医者 | 33 |
| 目 | 133 |
| 薬局 | 42 |
| 診察 | 18 |
| 説明 | 26 |

4.2 実験内容

比較実験は、2010年10月29日、30日に行った。実験協力者はTackPadの利用者である、情報系の大学生10名(男性10名)である。順序効果を考慮するため、実験協力者を各5名のグループA・Bに分けた。実験の流れは以下のとおりである。

- (1) 実験の説明を行った。
- (2) Aグループの協力者には本システム、Bグループの協力者には従来システムでの用例評価の方法を説明し、5分間、用例の評価活動を依頼した。
- (3) 利用するシステムを変えて、5分間、用例の評価活動を依頼した。
- (4) 両方のシステムでの評価活動が終わった後、アンケートを実施した。

比較実験では、評価する用例のタグを制限した。実験に利用したタグを表2に示す。これは、実験での評価時間が短いため、実験協力者がタグの選択に時間をとられないようにするためである。また、病名などの単語が少なく、評価しやすい用例群を評価対象とするためでもある。用例は日本語の用例のみ提示している。本システムおよび従来システムにおいて評価結果は共有されているため、どちらかのシステムで評価した用例はもう1つのシステムで評価することができない。

4.3 従来システムにおける評価活動

従来システムにおける用例の評価画面例を図8に示す。従来システムでの用例評価機能は「評価の負荷が少ない」ことを重点に設計されている。そのため、図8のように用例がリストで表示され、評価者は短時間に数多くの評価を行うことができる。

従来システムによる用例の評価の流れを以下に示す。

- (1) ユーザはタグの一覧から評価したいタグを選択する。
 - (2) 用例のリストとともに評価軸が表示される(図8)。ユーザは評価したい用例の評価軸上のラジオボタンをクリックして評価する。
- 実験で提示する項目を統一するために、この実験の間の



図 8 従来システムにおける用例の評価画面例

Fig. 8 Screenshot of evaluation scene on TackPad.

み、TackPad上でも1つの用例に対して「用例を評価した人数」を提示するようにした。

4.4 アンケートの結果

比較実験のアンケート結果を示し、考察を行う。アンケートでは、5段階のリッカートスケール(以下「5段階評価」と表記する)、または「従来システムと本システムのどちらがよいか」(以下「2件法」と表記する)を用いている。5段階評価では「1:強く同意しない」「2:同意しない」「3:どちらともいえない」「4:同意する」「5:強く同意する」の評価項目を使用した。また、有意差検定にはWilcoxonの符号付き順位検定を用い、有意水準5%で検定を行った。

4.4.1 達成感について

達成感に関するアンケートの結果を表3に示す。「システムを利用した用例の評価活動で達成感は得られたか」という質問を行ったところ(表3(1),(2))、従来システムでは中央値、最頻値ともに2、本システムでは中央値、最頻値ともに4という結果が得られた。これらの結果について有意差検定を行ったところ、有意確率0.005となり有意差が見られた。したがって、従来システムよりも本システムの方が達成感を与えられたと考えられる。自由記述では「(場所選択画面の)花が咲くアニメーションがよい」という意見が複数あり、用例群をいくつかに分けて提示することは、1つ1つに達成感を与えることができたと考えられる。また、「見た目に変化することで達成感を得られた」という意見もあり、用例を木として表示し、評価の様子を視覚化することによる効果もあったと考えられる。

4.4.2 評価状況の把握について

評価状況の把握に関するアンケートの結果を表4に示す。「システム全体でどれだけ用例に評価がつけられているかが分かりやすいか」という質問を行ったところ(表4(1),(2))、従来システムでは中央値1.5、最頻値1,2、本システムでは中央値、最頻値ともに4という結果が得られた。これらの結果について有意差検定を行ったところ、有意確率

表 3 達成感に関するアンケートの結果 (5段階評価)

Table 3 Results of the questionnaire survey on accomplishment (5-point Likert scale).

| | 質問項目 | 評価の分布 | | | | | 中央値 | 最頻値 |
|-----|------------------------------------|-------|---|---|---|---|-----|-----|
| | | 1 | 2 | 3 | 4 | 5 | | |
| (1) | 従来システムにて、多くの用例を評価することによって「達成感」を感じた | 3 | 5 | 2 | 0 | 0 | 2 | 2 |
| (2) | 本システムにて、多くの用例を評価することによって「達成感」を感じた | 0 | 0 | 2 | 6 | 2 | 4 | 4 |

・評価項目 (1: 強く同意しない 2: 同意しない 3: どちらともいえない 4: 同意する 5: 強く同意する)
 ・評価の分布の 1, 2, 3, 4, 5 の列は各評価値をつけた人数を示す。

表 4 評価状況の把握に関するアンケートの結果 (5段階評価)

Table 4 Results of the questionnaire survey on assessment of status (5-point Likert scale).

| | 質問項目 | 評価の分布 | | | | | 中央値 | 最頻値 |
|-----|-----------------------------------|-------|---|---|---|---|-----|------|
| | | 1 | 2 | 3 | 4 | 5 | | |
| (1) | 従来システム全体でどれだけ用例に評価がつけられているか分かりやすい | 5 | 5 | 0 | 0 | 0 | 1.5 | 1, 2 |
| (2) | 本システム全体でどれだけ用例に評価がつけられているか分かりやすい | 1 | 0 | 0 | 7 | 2 | 4 | 4 |

・評価項目 (1: 強く同意しない 2: 同意しない 3: どちらともいえない 4: 同意する 5: 強く同意する)
 ・評価の分布の 1, 2, 3, 4, 5 の列は各評価値をつけた人数を示す。

表 5 従来システムと本システムのどちらがよいか (2件法)

Table 5 Comparison between the old system and the new system (Select).

| | 質問項目 | 従来システム | 本システム |
|-----|------------------|--------|-------|
| (1) | 評価活動が楽しいのはどちらか | 0 | 10 |
| (2) | 評価活動をするのであればどちらか | 1 | 9 |

0.007 となり有意差が見られた。したがって、従来システムよりも本システムの方が評価状況の把握に役立つと考えられる。自由記述では「木の成長と関連しているので分かりやすい」という意見が複数あり、用例の評価状況を視覚化することで、評価状況の把握に役立つと考えられる。この結果から、用例評価の問題点である「評価すべき用例数が多く、ゴールが見えない」点を解決できたと考えられる。

4.4.3 従来システムと本システムとの比較

従来システムと本システムのどちらがよいか、というアンケートの結果を表 5 に示す。「評価活動が楽しいのはどちらか」という質問 (表 5(1)) については、実験協力者 10 名全員が「本システム」と回答しており、従来システムを利用するよりも楽しさを感じられるという結果が得られた。自由記述では「木が成長する姿を見るのは楽しい」という意見があり、用例評価を木の成長に見立てることにより、楽しさを与えることができたと考えられる。また、「評価活動をするのであればどちらか」という質問 (表 4(2)) では、実験協力者のうち 1 名が「従来システム」、9 名が「本システム」と回答した。「従来システム」と回答したユーザの自由記述には「作業の効率を考えたときは TackPad (従来システム) の方が楽」という意見があった。これらのこ

とから、多くの用例を手早く評価したいユーザは従来のシステムを、評価活動において楽しさや達成感を得たいユーザは本システムを利用するとよいと考えられる。

5. モチベーション維持効果検証のための実験

モチベーション維持効果を検証するための実験を行った。本実験の目的は、実験協力者が継続的に本システムを利用できるかどうかを検証することである。また、本実験よりシステムに追加した「あなたの森」「お知らせ」機能の効果の確認および、本システムを利用するユーザの行動調査を行う。

5.1 検証項目

本実験では、以下の 3 つの項目について検証を行う。

- (1) 実験協力者が本システムを継続的に利用できるかどうか。
- (2) 「あなたの森」機能および「お知らせ」機能の適用により、「自分の用例が評価されると、他のユーザの用例を評価する」効果があるか。
- (3) ユーザは本システムにおいてどのような行動をとるか。

表 6 本実験の日程

Table 6 Schedule of the experiment.

| 経過日 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|-------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|
| 実験の種類 | 募 | 募 | 募 | 公 | 公 | 公 | 公 | 公 | 公 | 公 | 公 | 公 | 公 | 公 |
| 告知メール | | | | ○ | | | | ○ | | | | | | |

・実験の種類 (募：利用者募集実験 公：一般公開実験)

5.2 実験内容

実験は、2010年12月13日から26日までの2週間行った。本実験は2段構成となっている。実験の日程を表6に示す。12月13日から15日までの3日間は利用者を募集した実験（以下、利用者募集実験と表記する）を行い、12月16日から26日までの11日間はシステムを一般公開し実験（以下、一般公開実験と表記する）を行った。これは利用者募集実験での実験協力者が、一般公開後もシステムを利用するかどうか検証するためである。表中の「告知メール」に関しては、5.2.2項において説明する。なお、利用者募集実験、一般公開実験ともに、本システムで提示される用例は、TackPadでの「理解できる言語」に依存する。たとえば、「理解できる言語」を「日本語、英語」と設定した実験協力者には、日本語と英語のタグおよび用例が提示される。

5.2.1 利用者募集実験

利用者募集実験では、TackPadの利用者にタスクを課して本システムを利用してもらった。利用者募集実験の協力者は、情報系の大学生10名（男性8名、女性2名）である。なお、協力者には利用者募集実験後に一般公開実験を行うことを伝えていない*5。

利用者募集実験の流れを以下に示す。タスクとして用例登録を課した理由は、「あなたの森」の変化を見てもらうためには、ユーザの森を作成する必要があるためである。

- (1) 実験の説明、システムの操作説明を行った。
- (2) 以下の作業を3日間毎日行ってもらった。
 - TackPadを利用して用例を5件以上登録する。
 - 本システムを利用して用例の評価活動を15分以上行う。

- (3) 3日目のタスクが終わった後、アンケートを実施した。

5.2.2 一般公開実験

一般公開実験では本システムを一般に公開し、利用者募集実験の協力者が引き続き本システムを利用するか検証している。実験の手順を以下に示す。各告知はTackPadメーリングリストと、TackPadのトップページに表示されるお知らせによって行っている。TackPadメーリングリストには、実験当時143名のユーザが登録されていた。アンケートは、期間中に「評価活動を行ったユーザ」および「お知

らせを確認したユーザ」に回答を依頼した*6。

- (1) 4日目に本システムを一般公開し、本システムの告知を行った*7。
- (2) 8日目に本システムに新しい機能を追加し、機能の追加について告知を行った*7,*8。
- (3) 14日目に、システムを利用したユーザにメールでアンケートを実施した。

5.3 利用者募集実験の結果と考察

利用者募集実験におけるアンケート結果と、システムログの分析結果を示し、考察を行う。アンケートでは、5段階評価、または「はい・いいえ」（以下「2件法」と表記する）を用いている。利用した機能に関する質問のみ回答を依頼したため、アンケート項目によって回答人数は異なる。

また、実験協力者のうち3名は用例評価のタスク（15分以上の評価活動）を達成できていない。1名は1日目の評価時間が13分、1名は1日目の評価時間が14分、1名は1日目から3日目の評価時間がそれぞれ13分、5分、7分であった*9。

5.3.1 あなたの森・お知らせ機能の効果について

「あなたの森」「お知らせ」機能に関するアンケートの結果を表7、表8に示す。また、ユーザが森を訪問した回数を表9に示す。「『あなたの森』の用例の木が成長したこと、他のユーザの用例にも評価したくなったか」という質問を行ったところ（表7）、中央値、最頻値ともに4となり、ユーザは自分が登録した用例が評価されることにより、他の用例を評価したいと思う傾向があるという結果が得られた。この理由として「恩返しとして評価したい」というコメントがあった。しかし、「『お知らせ』に提示されていたユーザの森に行ったか」「『お知らせ』に提示されていたユーザの用例を評価したか」という質問を行ったところ（表8(1), (2))、お知らせに提示されたユーザの森を訪問した実験協力者は1名であることが分かった。さらに、

*6 システムにアクセスしただけのユーザを除いて、アンケートの回答を依頼した。

*7 告知のメール、TackPadのお知らせの掲載は本実験4日目、8日目ともに朝11時から12時までの間に行った。

*8 「ユーザが登録した用例の木の『種類(見た目)』を変更する機能」を追加したが、今回の実験では新しい機能を追加したことによる効果は見られなかった。

*9 最後の1名は日本人であるが、TackPadの理解できる言語設定が英語のみに設定されており、この設定が影響した可能性がある。

*5 「近々システムを一般公開するので、今後ともよろしく願います」とだけ、実験終了後に伝えている。

表 7 アンケートの結果 (5 段階評価)

Table 7 Results of the questionnaire survey (5-point Likert scale).

| | 質問項目 | 評価の分布 | | | | | 回答 人数 | 中央 値 | 最頻 値 |
|-----|---|-------|---|---|---|---|----------|---------|---------|
| | | 1 | 2 | 3 | 4 | 5 | | | |
| (1) | 「あなたの森」の用例の木が成長していたことで、他のユーザの用例にも評価したくなった | 0 | 2 | 0 | 4 | 0 | 6 | 4 | 4 |

・評価項目 (1: 強く同意しない 2: 同意しない 3: どちらともいえない 4: 同意する 5: 強く同意する)
 ・評価の分布の 1, 2, 3, 4, 5 の列は各評価値をつけた人数を示す。

表 8 アンケートの結果 (2 件法)

Table 8 Results of the questionnaire survey (Select).

| | 質問項目 | 評価の分布 | |
|-----|-------------------------------|-------|-----|
| | | はい | いいえ |
| (1) | 「お知らせ」に提示されていたユーザの森に行きましたか? | 1 | 4 |
| (2) | 「お知らせ」に提示されていたユーザの用例を評価しましたか? | 1 | 4 |

・回答人数は 5 名である。

表 9 各ユーザが森を訪問した回数

Table 9 Visit frequency of "Forest" for each user.

| | 森の種類 | A | B | C | D | E | F | G | H | I | J |
|-----|-----------------|----|----|----|----|----|----|----|----|----|----|
| (1) | 用例についたタグで分類された森 | 1 | 21 | 42 | 5 | 21 | 25 | 35 | 40 | 50 | 22 |
| (2) | 登録したユーザで分類された森 | 22 | 8 | 8 | 42 | 6 | 0 | 1 | 0 | 4 | 0 |

・A~J はユーザを表す。
 ・別の森へ移動した回数をカウントしている。

表 10 訪問したい森を選択する基準

Table 10 Criteria for user's selection of "Forest."

| | 森を選択するときに参考にした項目 | 人数 |
|-----|------------------|----|
| (1) | 小さな森 (用例の少ない森) | 6 |
| (2) | 大きな森 (用例の多い森) | 3 |
| (3) | 行きたい森リストの上部の森 | 2 |
| (4) | 知り合いのユーザの森 | 2 |

・回答人数は 10 名である。
 ・この質問の回答形式は複数選択である。

表 9 より実験協力者が訪問する森は、タグで分類された森、ユーザで分類された森のどちらかに偏っていたことが分かる。すなわち、特定の実験協力者はユーザで分類された森をあまり利用しなかった。これらのことから、評価してくれたユーザの森へ手軽に訪問できる機能が必要であると考えられる。

5.3.2 訪問したい森の基準

「どんなことを基準に訪問したい森を選択するか」という質問を行ったところ (表 10)、用例が少ない小さな森を訪問するユーザが多いことが分かった。この理由として、「すべての用例に評価をつけることが容易であり、手軽に達成感を得ることができる」という意見があった。また、逆に用例の多い「大きな森」を基準にする実験協力者からは、「用例が多いほうが評価したときに達成感があったため」というコメントがあった。「小さな森」を評価することで得られる手軽な達成感、「大きな森」を評価することで得られる大きな達成感、それぞれを増幅させる仕組みがあれ

ばよいのではないかと考えられる。

5.4 一般公開実験の結果と考察

一般公開実験のアンケート結果と、システムログの分析結果を示し、考察を行う。アンケートでは、5段階評価、または「はい・いいえ」(以下「2 件法」と表記する)を用いている。アンケートは 20 代のユーザ 5 名から回答を得た。5 名のうち 3 名は、利用者募集実験の協力者である。また、アンケートの回答者のうち 3 名は評価活動を行っており、残りの 2 名は「お知らせ」を確認するのみであった。

一般公開実験の期間中にシステムにアクセスしたユーザの情報を表 11 に、実験期間中のアクセス回数を表 12 に示す。表 12 中の「アクセス回数」は、ユーザの直前のアクセス時の行動から 30 分以上たったあとのアクセスを別のアクセスとし、カウントしている。この 30 分という時間は、Google Analytics などのアクセス解析システムで利用されている時間 [14] と同じである。また、「アクセスのみ

表 11 実験期間中にアクセスしたユーザの分類
Table 11 Kinds of user's access on term of the experiment.

| ユーザ | A* | B* | C | D | E | F | G | H |
|-------------|----|----|---|---|---|---|---|---|
| 利用者募集実験の協力者 | ○ | ○ | ○ | - | - | ○ | ○ | ○ |
| アンケートの回答者 | ○ | ○ | ○ | ○ | ○ | - | - | - |
| アクセスのみ | - | - | - | - | - | ○ | ○ | ○ |

*: ユーザ A, B は評価活動を行っておらず, お知らせを確認するのみであった.

表 12 一般公開実験の期間中のユーザのアクセス回数
Table 12 Number of user's access on term of open to the public experiment.

| ユーザ/経過日 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---------|---|---|---|---|---|---|----|----|----|----|----|
| A | 1 | | | | | | | | | | |
| B | 1 | | | | | | | | | | |
| C | | | | | 1 | | | | | | |
| D | | | | | 1 | | | | | | |
| E | | | | | | | | | | | 1 |
| F | | | | | | | | 1 | | | |
| G | | 1 | | | | | | | | | |
| H | 1 | | | | | | | | | | |
| アクセスの合計 | 3 | 1 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 1 |

- ・表中の数字はアクセス回数を示している.
- ・4日目と8日目にメーリングリストにより告知を行っている.
- ・アクセス回数は, ユーザの前回のアクセスの行動から30分以上たったあとのアクセスを別アクセスとしてカウントしている [14].

表 13 5段階評価のアンケートの結果
Table 13 Results of the 5-point-Likert-scale questionnaire survey.

| | 質問項目 | ユーザの評価 | | | | | 中央値 | 最頻値 |
|-----|----------------------|--------|---|---|---|---|-----|-----|
| | | A | B | C | D | E | | |
| (1) | 今後もシステムを利用して用例を評価したい | 4 | 4 | 3 | 2 | 4 | 4 | 4 |

- ・評価項目 (1: 強く同意しない 2: 同意しない 3: どちらともいえない 4: 同意する 5: 強く同意する)
- ・ユーザの評価の各数値は, 評価値を示す.
- ・A, B, C のユーザは利用者募集実験の参加者である.

のユーザ」の中には「画面遷移を行ったのみユーザ」(「様子見」のユーザ) も含まれる.

一般公開実験の期間中にはメーリングリストによる告知を受け取る 143 名のうち, 8 名が本システムにアクセスした. その中で評価活動を行ったユーザが 3 名, お知らせのみを確認したユーザが 2 名, アクセスのみのユーザは 3 名である. 利用者募集実験に参加したユーザは, 10 名中 6 名が本システムにアクセスした. そのうち, 評価活動を行ったユーザは 1 名, お知らせを確認したユーザが 2 名, アクセスのみのユーザが 3 名である. 表 12 より, どのユーザも期間中 1 度しかアクセスしておらず, 継続的に評価活動を行っていない. しかし, 本実験開始前の 2 週間において, TackPad における用例評価機能を用いて評価を行ったユーザは 2 名で, 各アクセスは 1 度であり, これと比較して本システムにアクセスしたユーザは多い. また, 表 13 で示す, 一般公開実験後のアンケート結果より, 質問「今後もシステムを利用して用例を評価したい」に対する利用

者募集実験参加者 A, B, C の回答は, 5 段階評価で 3, 4, 4 となっている. 本実験でアクセスしたユーザの中で利用者募集実験に参加したユーザが多いこと, TackPad における用例評価機能を利用したユーザとの比較, アンケートの結果より, 本システムはユーザのモチベーション維持支援の可能性があると考えられる. 十分なモチベーション維持支援のためには, ユーザが継続してシステムを利用できる「きっかけ」が必要であると考えられる. また, 今回 2 週間という短い期間での検証であったため, より長期的な検証が必要であると考えられる.

6. おわりに

本論文では, 「楽しさ」や「達成感」などの要素を用いた, 用例評価活動のモチベーション維持支援システム「用例の森」を開発した. 本システムの有用性を示すために, 従来システムでの評価活動との比較実験, および評価活動を継続できるかどうかを検証するために実験を行った. 本

論文の貢献は以下の4点にまとめられる。

- (1) 非日常的な作業に対するモチベーション維持支援システム「用例の森」の提案を行い、実現した。「用例の森」は、非日常的な作業である用例の評価活動において、モチベーション維持支援の可能性があることを示した。
- (2) 用例を木に、用例の評価状況を木の成長に見立てることにより、ユーザに「楽しさ」や「達成感」を与えることができ、「評価状況の把握」に役立つことが分かった。また、用例群をいくつかのグループに分け、小さな目標としてユーザに提示することにより、達成感を与えることの補助ができることが分かった。
- (3) ユーザは自分の登録した用例が評価されると、他のユーザが登録した用例を評価したいと思う傾向があることを明らかにした。
- (4) ユーザは手軽に達成感を得るために、数の少ない用例群を優先して評価する傾向があることを明らかにした。

今後の課題としては、インタフェースの改善、他のユーザが自分の用例を評価したことを効果的に提示、ユーザが継続してシステムを利用できる「きっかけ」の提供があげられる。これらの改善を行い、長期的なモチベーション維持効果を検証する。

謝辞 本研究の一部は、総務省戦略的情報通信研究開発推進制度 (SCOPE) の平成 22 年度採択課題「医療現場における利用者適応型多言語間コミュニケーション支援のための基盤技術の研究開発」の補助を受けた。

参考文献

- [1] 金谷裕幸, 小林智也, 千葉慶人, 伊藤直樹, 西本一志: モーニング・コミュニティ: 起床時のモチベーションを強化する社会的目覚まし時計の提案, *インタラクション* 2007, pp.221-222 (2007).
- [2] Fujiki, Y., Kazakos, K., Puri, C., Buddharaju, P., Pavlidis, I. and Levine, J.: NEAT-o-Games: Blending physical activity and fun in the daily routine, *Computers in Entertainment (CIE)*, Vol.6, No.2, pp.1-22 (2008).
- [3] 倉本 到, 芝田祐也, 島田卓哉, 渋谷 雄, 辻野嘉宏: 日常生活の日常生活による日常生活のためのエンタテインメント, *エンタテインメントコンピューティング 2006 講演論文集*, pp.117-118 (2006).
- [4] Sánche, J. and Salinas, A.: Science problem solving learning through mobile gaming, *Proc. 12th International Conference on Entertainment and Media in the Ubiquitous Era*, pp.49-53, ACM (2008).
- [5] 法務省: 2009 年度の結果の概要, 登録外国人統計, 法務省 (オンライン), 入手先 (<http://www.moj.go.jp/content/000052445.pdf>) (参照 2011-10-08).
- [6] 田村太郎: 多民族共生社会ニッポンとボランティア活動, 明石書店 (2000).
- [7] 福島 拓, 宮部真衣, 吉野 孝, 重野亜久里: 医療分野を対象とした多言語用例対訳収集 Web システム Tack-Pad の開発, マルチメディア, 分散, 協調とモバイル (DICOMO2008) シンポジウム, pp.1030-1036 (2008).
- [8] 岩部正明, 村上陽平, 重野亜久里, 石田 亨: Web サー

ビス連携を用いた医療用例対訳の収集と利用, 電子情報通信学会技術研究報告, AI2006-28, pp.17-22 (2006).

- [9] Sonnenfeld, J.: Commentary: Academic Learning, Worker Learning, and the Hawthorne Studies, *Social Forces*, Vol.61, p.904 (1982).
- [10] 倉本 到, 片山拓馬, 渋谷 雄, 辻野嘉宏: 懐優館: 作業意欲を持続的に維持向上させる EELF に基づく主観的比較型エンタテインメントシステム, *情報処理学会論文誌*, Vol.50, No.12, pp.2807-2818 (2009).
- [11] 山田和明, 中小路久美代, 山本恭裕: オンラインコミュニティにおける知識共創のモデル, 人工知能学会, 第四回知識流通ネットワーク研究会, 人工知能学会 (2009).
- [12] Kuznetsov, S.: Motivations of contributors to Wikipedia, *ACM SIGCAS Computers and Society*, Vol.36, No.2, article 1 (2006).
- [13] Pink, D.H. (著), 大前研一 (訳): モチベーション 3.0 (原題: Drive), 講談社 (2010).
- [14] Google: Google セッション - Analytics ヘルプ, Google (オンライン), 入手先 (<http://www.google.com/support/analytics/bin/answer.py?hl=ja&answer=33073>) (参照 2011-10-08).



狩野 翔 (学生会員)

昭和 63 年生。平成 23 年和歌山大学システム工学部デザイン情報学科卒業。現在、同大学大学院システム工学研究科システム工学専攻博士前期課程在学中。モチベーション維持支援に関する研究に従事。



福島 拓 (学生会員)

昭和 61 年生。平成 20 年和歌山大学システム工学部デザイン情報学科中退。平成 22 年同大学大学院システム工学研究科システム工学専攻博士前期課程修了。現在、同大学院システム工学研究科システム工学専攻博士後期課程在学中。多言語間コミュニケーション支援に関する研究に従事。



吉野 孝 (正会員)

昭和 44 年生。平成 4 年鹿児島大学工学部電子工学科卒業。平成 6 年同大学大学院工学研究科電気工学専攻修士課程修了。現在、和歌山大学システム工学部デザイン情報学科准教授。博士 (情報科学)。コミュニケーション支援の研究に従事。