

令和 6 年 5 月 30 日現在

機関番号：32644

研究種目：基盤研究(C)（一般）

研究期間：2020～2023

課題番号：20K12411

研究課題名（和文）SNS収集データをもとにしたAI技術を用いたWeb情報の推薦

研究課題名（英文）Recommendation of web information using AI technology based on SNS collected data

研究代表者

星野 祐子（Hoshino, Yuko）

東海大学・情報通信学部・講師

研究者番号：80435271

交付決定額（研究期間全体）：（直接経費） 2,700,000円

研究成果の概要（和文）：本研究は、ユーザがWebページを閲覧している間の視線情報を用いて注視箇所を特定し、その箇所の文章や画像を抽出、ユーザの嗜好を把握し、パーソナライズされた検索キーワードや観光関連Webニュースを推薦するシステムを開発した。システムは文章・画像抽出プロセス、嗜好把握プロセス、関連情報提示プロセスの3つから構成される。実験では、観光スポットに関する推薦情報の適合度が高いことが示された。今後は、文化的背景の異なるグループ間の比較や読みの深さを考慮した嗜好抽出精度の向上が課題とされる。

研究成果の学術的意義や社会的意義

この研究は、視線追跡技術と機械学習を組み合わせることでユーザの興味をリアルタイムで特定し、パーソナライズド推薦を行う手法を開発しました。学術的には、ユーザ行動解析と多言語対応の情報推薦システムにおける新しいアプローチを提供します。社会的には、旅行者の観光体験を向上させ、広告やマーケティングの効果を高め、多文化理解を促進します。これにより、技術革新と社会的影響の両面で貢献できるものと考えます。

研究成果の概要（英文）：This study developed a system that utilizes gaze information while users are browsing web pages to identify areas of focus. The system extracts text and images from these areas, understands user preferences, and recommends personalized search keywords and tourism-related web news. The system consists of three processes: text and image extraction, preference understanding, and relevant information presentation. Experiments showed a high degree of accuracy in the recommendation of information related to tourist spots. Future challenges include comparing groups with different cultural backgrounds and improving the accuracy of preference extraction by considering the depth of reading.

研究分野：ウェブ情報学

キーワード：Web 情報推薦 視線情報 観光情報 機械学習 ヒューマンインタラクション

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

様式 C - 19 , F - 19 - 1 (共通)

1. 研究開始当初の背景

(1) 観光客が情報収集するためにインターネットが利用されている。中でも FaceBook や Twitter などの SNS が役に立ったという意見が目立っている。しかしながら、これら SNS は小さな情報が広く分散して存在しており、ユーザが欲している情報を手に入れるには手間がかかる。SNS で提供されている情報は古いものや正しくないもの、ユーザが本来求めている情報とは異なるものも多く存在する。ユーザは自分で正しい情報、必要な情報を選択しなければならない。また、20 代の若者は SNS を情報収集の手段として利用している人が多いが、少し上の世代になると情報収集手段として SNS を活用できている人が少ない。

(2) 多くの Web ページで発信されている観光情報は日本語、英語、その他外国語共にほぼ同じ内容であるが、様々な国からの訪日客が同じものに興味を持つとは限らない。よって、提供される内容も観光客の出身地域などによる文化的背景を考慮すべきである。しかしながら、これを実現することは Web 技術者やサーバ管理者への負荷が大きく、かつ異文化を理解する必要もあり難しいものである。これらが本研究の背景である。

2. 研究の目的

本研究の目的は SNS から収集した日本語と英語テキストデータに画像データを加え AI による学習を用いることで、観光客それぞれが興味を持つスポットやイベントの違いをより明確に抽出することである。加えて、Web 閲覧履歴等を組み合わせることで、使用言語の違いをもとにした文化的背景を考慮した話題抽出だけでなく、より個人の好みに合わせた、より必要とされる話題、情報の抽出と提供を実現することである。しかし、テキスト、画像の両収集データから趣向やユーザの欲する情報分野を特定する方法論やモデルは存在していない。

そこで、本研究では観光情報 Web ページや投稿された SNS データを収集し、言語別に機械学習を用いたテキストデータ群からの潜在的トピック抽出、ディープラーニングを用いた収集画像から新たに話題となり始めたスポット等の抽出を行い、観光情報 Web ページ等への掲載推奨情報等への掲載推奨情報として推薦する。加えて、Web 閲覧履歴から抽出した個々人の趣向を取り込むことで、よりその人にマッチした Web 掲載情報を推薦、提示できるようにし、SNS を情報収集手段として活用することが難しい人でも必要な情報を得ることができるようにする。

3. 研究の方法

まず、図 1 のようにユーザがブラウザを通じて SNS を含む Web ページを閲覧している間、ユーザの視線情報を活用して Web ページ上で特に注視していた場所を特定する手法を開発した。

次に特定した場所に掲載されていた文章および画像を抽出し、抽出した情報をもとに機械学習、深層学習等の技術も活用してユーザの嗜好を把握し、よりパーソナライズされた Web 検索キーワード、観光関連 Web ニュースの推薦を行う観光情報推薦システムを開発した。

図 2 に示すように、このシステムは主に 3 つの部分からなり、文章・画像抽出プロセス、嗜好把握プロセス、関連情報提示プロセスで構成される。文章・画像抽出プロセスでは、ユーザの視線運動（読み速度）を計測し、熟読（興味あり）と判断された部分の文章を HTML 中の p タグの位置をもとに取得して、ページ遷移（ブラウザの戻るボタンを押下等）したときに、取得した文章を要約し、ユーザプロフィールとして保存する。取得したエリア内に画像があった場合には画像データも取得し、画像に含まれるオブジェクトやその状態を AI 技術によりテキスト化し、Web ページから取得した文章と合わせている。ユーザが一連の Web 閲覧行動を終了し、システム上の「suggest」ボタンを押したとき、ユーザプロフィール

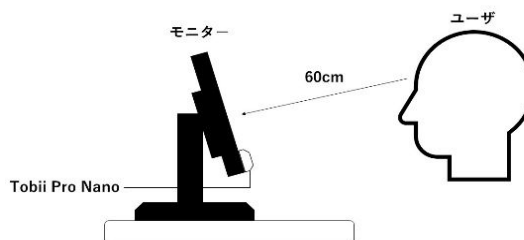


図 1 Web ページ閲覧時の環境

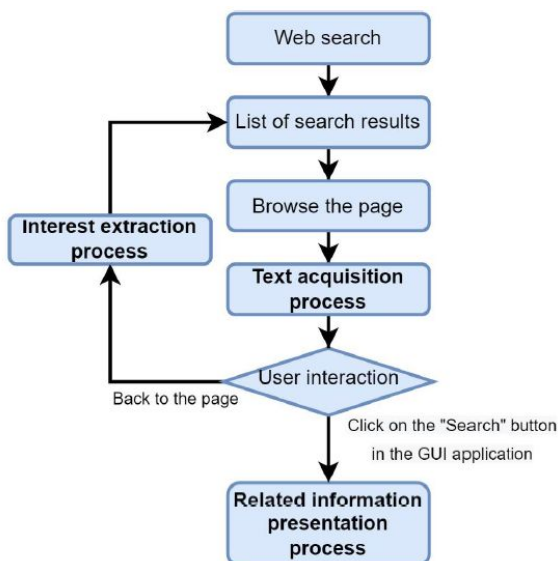


図 2 本システムの処理の流れ

と最後に閲覧していたページの注視情報からユーザ嗜好を抽出する。嗜好の抽出には保存された要約文と最後に抽出された注視文章とのコサイン類似度を用いている。関連連情報提示プロセスでは TF-IDF による重要語抽出から得られた語と関連度の高い語、共起確率の高い語、閲読文章に含まれる単語の感情分析結果と観光スポットの口コミ情報 (Web から事前にユーザが閲覧している観光地周辺のスポット情報を取得) に含まれる単語の感情分析結果が類似しているものの中から適合度の高い上位 5 件ずつを提示している。観光関連の Web ニュース提示では口コミ情報ではなく Web ニュースからデータを取得している。

最後に参加者を募り、システムの評価を行った。この実験では東京近郊の観光スポットについて「一般的な周辺観光」に興味がある人、「歴史観光」に興味がある人を想定して参加者に Web 閲覧をしてもらい、推薦された情報および適合度に関する実験後アンケートにより評価を行った。なお、「一般的な周辺観光に興味がある」または「歴史観光興味がある」のグループ分けは参加者に周辺観光に関する問題で構成された小テストまたは歴史刊行に関する問題で構成された小テストを解いてもらうことで行った。検証時に参加者が閲覧した Web ページは「smart magazine (<https://www.smartmagazine.jp/tokyo/>)」である。

また、2022 年度までのシステムでは日本語文章に特化したシステムであったが、API の終了もあり、2023 年度に画像のテキスト化および要約文の生成を行う部分を英語にも対応できるもの (GPT4-vision-preview, GPT3.5Turbo) に入れ替えている。

4. 研究成果

本研究で開発したシステムの実行画面の例を図 3 に示す。本システムでは作成した GUI メニュー画面上の Gaze Mode を選択することで、chrome ブラウザと視線検出による HTML 要素の取得を行うプログラムを起動する。Gaze Mode を終了するまで、本システムはユーザの注視位置を取得し続け、Web ブラウザ上でユーザが注視していると判断された部分の<p>タグ内のコンテンツを取得する。

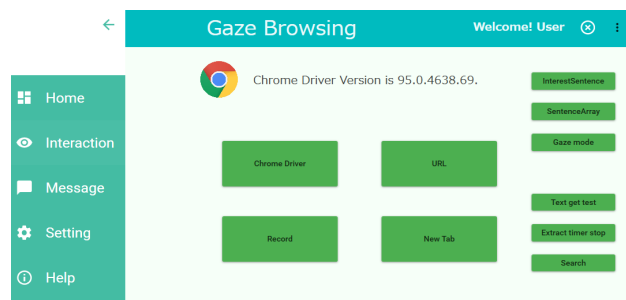


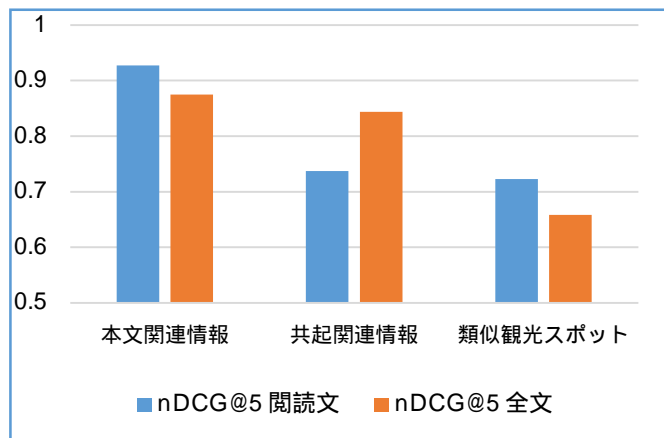
図 3 メニュー画面の GUI

実験参加者が観光情報検索キーワードを推薦するシステムを使用し、smart magazine の東京・神宮外苑に関する Web ページを閲覧した時の実行結果例を図 4 に示す。参加者を「一般的な周辺観光」に興味がある人グループ、「歴史観光」に興味がある人グループに分け推薦アイテム (上位 5 個) とアンケートによるデータをもとに (normalized Discounted Cumulative Gain) 及び MAP (Mean Average Precision) による検証を行った。



図 4 キーワード推薦実行結果例

検証の結果、「周辺観光」グループの本文関連情報および類似観光スポットの推薦において閲覧した Web ページ全文を対象とした推薦処理に比べ、本手法による推薦処理での適合度が高くなり、推薦精度も高いことが示唆された。「歴史観光」グループでは今回の検証で使用した Web ページに掲載されていた観光スポットが歴史的要素を多く含む場所であったこともあり、本文関連情報での適合度に差は認められなかった。また、「類似観光スポット」に関しては Web ページ全文を対象とした推薦処理の適合度が高くなった。「周辺観光」グループでの全文を対象とした推薦処理とユーザ閲読文を対象とした推薦処理による適合度を図 5 に示す。



閲読文 (本研究の手法) : 被験者に Web 閲覧結果から推薦

全文 : Web ページ全文を解析した結果から推薦

図 5 周辺観光の適合度

加えて表1に示すようにウィルコクソンの符号順位和による有意差検定を行った結果、本文関連情報ではWebページ全文を対象とした推薦処理と本研究のユーザ読文を対象とした推薦処理での有意差が認められた。従って、本手法によるWeb閲覧情報から嗜好・興味を考慮した推薦が可能であることが示唆されたと考える。

今回の手法では個人ごとの検索意図、嗜好を考慮した推薦ができることが示唆されているが、日本人と外国人など文化的背景が異なるグループ同士での比較ができていない。今後は類似した文化的背景を持つユーザグループでの推薦比較を行うことを考えている。また、ユーザ嗜好の抽出精度を上げるため、読み直しの有無など「どのように読んでいるのか」といった読みの深さも扱うようにしたい。

表1 ウィルコクソンの符号順位和検定による

	A群(周辺観光)	B群(歴史観光)
本文関連情報	0.03292 *	0.75961
共起関連情報	0.2703	0.36269
類似観光スポット	0.65413	0.34545

帰無仮説 H_0 : 推薦精度に差がない

対立仮説 H_1 : 推薦精度に差がある

5. 主な発表論文等

〔雑誌論文〕 計8件（うち査読付論文 8件/うち国際共著 0件/うちオープンアクセス 5件）

1. 著者名 Mori Taiga, Hoshino Yuko, Yamada Mitsuho, Ishii Eriko	4. 巻 15
2. 論文標題 Development of a Tourist Information Search Support System that Reflects Interests Based on User's Gaze	5. 発行年 2022年
3. 雑誌名 International Conference on Tourism Research	6. 最初と最後の頁 562～567
掲載論文のDOI（デジタルオブジェクト識別子） 10.34190/ictr.15.1.152	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 森大河、山田光穂、石井英里子、星野祐子	4. 巻 16
2. 論文標題 視線情報を活用したWeb検索支援システムの開発	5. 発行年 2021年
3. 雑誌名 パーソナルコンピュータ利用技術学会論文誌	6. 最初と最後の頁 20～29
掲載論文のDOI（デジタルオブジェクト識別子） 10.24663/jjpcats.16.2_20	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 野坂祐介、石井英里子、星野祐子、山田光穂	4. 巻 16
2. 論文標題 鑑賞時間を变化させたときの絵画鑑賞中の注視点の比較	5. 発行年 2021年
3. 雑誌名 パーソナルコンピュータ利用技術学会論文誌	6. 最初と最後の頁 30～39
掲載論文のDOI（デジタルオブジェクト識別子） 10.24663/jjpcats.16.2_30	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 篠原未歩、石井英里子、星野祐子、山田光穂	4. 巻 16
2. 論文標題 高品質な映像条件が人にもたらす癒やし効果の検討	5. 発行年 2021年
3. 雑誌名 パーソナルコンピュータ利用技術学会論文誌	6. 最初と最後の頁 40～55
掲載論文のDOI（デジタルオブジェクト識別子） 10.24663/jjpcats.16.2_40	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Nosaka Yusuke, Ishii Eriko, Yamada Mitsuho, Hoshino Yuko	4. 巻 1
2. 論文標題 Proposal for a saliency map based on the focal points of a painting	5. 発行年 2021年
3. 雑誌名 10th Global Conference on Consumer Electronics	6. 最初と最後の頁 982-985
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE53005.2021.9621824	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Shinohara Miho, Ishii Eriko, Yamada Mitsuho, Hoshino Yuko	4. 巻 1
2. 論文標題 Comparison of Healing Effect on Biomedical Signals by the Difference of Color Gamut and Dynamic Range When Viewing Nature Images	5. 発行年 2021年
3. 雑誌名 10th Global Conference on Consumer Electronics	6. 最初と最後の頁 666-669
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE53005.2021.9621971	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Nosaka Yusuke, Shinohara Miho, Nomura Kosuke, Sarugaku Takuya, Yamada Mitsuho	4. 巻 6
2. 論文標題 Analysis of Gaze Time Spent at the Gazing Point that is Required During Reading	5. 発行年 2021年
3. 雑誌名 Advances in Science, Technology and Engineering Systems Journal	6. 最初と最後の頁 726 ~ 734
掲載論文のDOI (デジタルオブジェクト識別子) 10.25046/aj060179	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 野坂 祐介、猿樂 拓也、石井 英里子、星野 祐子、山田 光穂	4. 巻 15
2. 論文標題 絵画鑑賞時の注視点分布のデータベース作成	5. 発行年 2021年
3. 雑誌名 パーソナルコンピュータ利用技術学会論文誌	6. 最初と最後の頁 21 ~ 31
掲載論文のDOI (デジタルオブジェクト識別子) 10.24663/jjpcats.15.1_21	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計28件（うち招待講演 0件 / うち国際学会 9件）

1. 発表者名 森 大河 山田 光穂 石井 英里子 星野 祐子
2. 発表標題 Web 閲覧中のユーザの視線に基づく関連情報提示システムの開発及び評価
3. 学会等名 FIT2022(第21回情報科学技術フォーラム)
4. 発表年 2022年

1. 発表者名 森 大河, 山田光穂, 石井英里子, 星野祐子
2. 発表標題 視線情報の活用による観光情報推薦システムの開発及び評価
3. 学会等名 電子情報通信学会(IMQ, IE, MVE, CQ) 合同研究会
4. 発表年 2023年

1. 発表者名 浅井杜和, 石井英里子, 島立短大, 山田光穂, 星野祐子
2. 発表標題 写真の代表色抽出手法の比較とオブジェクト抽出手法の調査
3. 学会等名 電子情報通信学会(IMQ, IE, MVE, CQ) 合同研究会
4. 発表年 2023年

1. 発表者名 藤坂直輝, 石井英里子, 山田光穂, 星野祐子
2. 発表標題 Web上で動作するプログラミング学習支援システムの研究
3. 学会等名 電子情報通信学会(IMQ, IE, MVE, CQ) 合同研究会
4. 発表年 2023年

1. 発表者名 滝健太郎, 山田光穂, 石井英里子, 星野祐子
2. 発表標題 口コミ分析による複数の観光要素を考慮した観光スポット推薦システム
3. 学会等名 第7回 国際ICT利用研究学会 全国大会
4. 発表年 2022年

1. 発表者名 森大河, 山田光穂, 石井英里子, 星野祐子
2. 発表標題 視線情報の活用による検索キーワード推薦システムの開発
3. 学会等名 第7回 国際ICT利用研究学会 全国大会
4. 発表年 2022年

1. 発表者名 森 大河, 山田, 光穂, 石井, 英里子, 星野 祐子
2. 発表標題 Web閲覧中の視線情報を用いた観光情報推薦システムの開発
3. 学会等名 観光情報学会 第18回全国大会
4. 発表年 2022年

1. 発表者名 森大河, 大友隆秀, 野坂祐介, 石井英里子, 星野祐子, 山田光穂
2. 発表標題 視線情報の活用によるWeb検索支援システムの開発
3. 学会等名 第14回データ工学と情報マネジメントに関するフォーラム DEIM2022
4. 発表年 2022年

1. 発表者名 Shimul Rakibul Hasan, Eriko Ishii, Mitsuho Yamada, Yuko Hoshino
2. 発表標題 Extracting Sightseeing-Related Information from Social Networking Service
3. 学会等名 第14回データ工学と情報マネジメントに関するフォーラム DEIM2022
4. 発表年 2022年

1. 発表者名 Taiga Mori, Mitsuho Yamada, Eriko Ishii, Yuko Hoshino
2. 発表標題 Development and Evaluation of a System for Presenting Information Related to Web Pages Based on the User's Gaze
3. 学会等名 IDW '21 - 28th International Display Workshops (国際学会)
4. 発表年 2021年

1. 発表者名 森 大河, 山田 光穂, 石井 英里子, 星野 祐子
2. 発表標題 Web閲覧時の視線情報を活用した興味語推薦システムの開発
3. 学会等名 映像情報メディア学会 2021年冬季大会
4. 発表年 2021年

1. 発表者名 森 大河, 山田 光穂, 石井 英里子, 星野 祐子
2. 発表標題 Web 閲覧中の視線情報を活用した検索キーワード推薦システムの開発
3. 学会等名 第16回パーソナルコンピュータ利用技術学会 全国大会
4. 発表年 2021年

1. 発表者名 Shimul Rakibul Hasan, Eriko Ishii, Mitsuho Yamada, Yuko Hoshino
2. 発表標題 Extracting Sightseeing-Related Information from Social Networking Service
3. 学会等名 第20回情報科学技術フォーラム(FIT2021)
4. 発表年 2021年

1. 発表者名 Shimul Rakibul Hasan, Eriko Ishii, Mitsuho Yamada, Yuko Hoshino
2. 発表標題 Extracting Sightseeing Related Information from Social Networking Service Data
3. 学会等名 4th International Conference on Tourism Research 2021 (国際学会)
4. 発表年 2021年

1. 発表者名 Taiga Mori, Mitsuho Yamada, Eriko Ishii, Yuko Hoshino
2. 発表標題 Development of a Related Tourism Information Provision System Using Traveler's Gaze Information
3. 学会等名 4th International Conference on Tourism Research 2021 (国際学会)
4. 発表年 2021年

1. 発表者名 Shimul Rakibul Hasan, Eriko Ishii, Mitsuho Yamada, Yuko Hoshino
2. 発表標題 Extracting Information from Social Networking Service about Sightseeing Locations
3. 学会等名 IDW '21 - 28th International Display Workshops (国際学会)
4. 発表年 2021年

1. 発表者名 Yuko Hoshino, Eriko Ishii, Yuta Nakamura, Takahide Ohtomo, Mitsuho Yamada
2. 発表標題 GATHERING IMAGE DATA FROM SNS AND TRAINING MACHINE LEARNING MODELS FOR IMAGE CLASSIFICATION
3. 学会等名 The Tenth International Workshop on Image Media Quality and its Applications, IMQA2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Yusuke Nosaka, Yuko Hoshino, Mitsuho Yamada
2. 発表標題 Analysis of Gaze Points When Looking at Paintings and Saliency Map to Improve the Accuracy of ROI (Region of Interest)
3. 学会等名 8th Conference on Embedded Networked Sensor Systems (国際学会)
4. 発表年 2020年

1. 発表者名 野坂祐介, 猿樂拓也, 星野 祐子, 山田 光穂
2. 発表標題 絵画鑑賞時の注視点分布と絵画のフォーカルポイントとの比較
3. 学会等名 FIT2020 (第19 回情報科学技術フォーラム)
4. 発表年 2020年

1. 発表者名 野坂祐介, 石井英里子, 星野 祐子, 山田 光穂
2. 発表標題 絵画鑑賞時の注視点データベースの提案と顕著性マップとの比較
3. 学会等名 信学技報, CQ2020-111
4. 発表年 2020年

1. 発表者名 野坂祐介, 猿樂拓也, 星野 祐子, 山田 光穂
2. 発表標題 絵画のフォーカルポイントに基づいた絵画鑑賞時の注視点分布と顕著性マップの一致度についての考察
3. 学会等名 映像情報メディア学会創立70周年記念大会
4. 発表年 2020年

1. 発表者名 シムル ロキブル ハサン, 星野祐子, 山田 光穂
2. 発表標題 SNS 画像からのオブジェクト検出と観光スポットに関する 投稿リスト作成システムの提案
3. 学会等名 第 15 回パーソナルコンピュータ利用技術学会全国大会
4. 発表年 2020年

1. 発表者名 Shu Iwasa, Taiga Mori, Eriko Ishii, Mitsuho Yamada, Yuko Hoshino
2. 発表標題 A Research on Keyword Recommendation via Eye Gaze when Browsing the Web
3. 学会等名 1st the International Conference on ICT application research (IAR) (国際学会)
4. 発表年 2023年

1. 発表者名 Towa Asai, Eriko Ishii, Mitsuho Yamada, Yuko Hoshino
2. 発表標題 Comparing Extraction Methods for Landscape Images Analysis
3. 学会等名 1st the International Conference on ICT application research (IAR) (国際学会)
4. 発表年 2023年

1. 発表者名 浅井 杜和, 石井 英里子, 山田 光穂, 星野 祐子
2. 発表標題 景観画像に対する代表色抽出手法の比較とオブジェクト抽出の精度調査
3. 学会等名 FIT2023 (第 22 回情報科学技術フォーラム)
4. 発表年 2023年

1. 発表者名 岩佐周, 森大河, 星野祐子, 石井英里子, 山田光穂
2. 発表標題 Web閲覧時の視線情報を取り入れたニュース検索ワード推薦システムの提案
3. 学会等名 第8回 国際ICT利用研究学会全国大会
4. 発表年 2023年

1. 発表者名 Shu Iwasa, Taiga Mori, Eriko Ishii, Mitsuho Yamada, Yuko Hoshino
2. 発表標題 PROPOSAL FOR A RELATED NEWS RECOMMENDATION METHOD USING THE GAZE WHILE VIEWING NEWS ARTICLES
3. 学会等名 The Twelfth International Workshop on Image Media Quality and Its Applications, IMQA2024 (国際学会)
4. 発表年 2024年

1. 発表者名 星野 祐子, 柳川 凜太郎, 清田 真一, 岩佐 周, 石井 英里子, 山田 光穂
2. 発表標題 視線を活用したWeb 閲覧データに基づくパーソナライズドニュース推薦システムに関する研究
3. 学会等名 電子情報通信学会 信学技報 IMQ研究会
4. 発表年 2024年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担者	山田 光穂 (Yamada Mitsuho) (60366086)	東海大学・情報通信学部・教授 (32644)	
研究 分担者	石井 英里子 (Ishii Eriko) (80580878)	鹿児島県立短期大学・文学科 英語英文学専攻・准教授 (47701)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------