

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 6 月 1 日現在

機関番号：62615

研究種目：基盤研究（B）

研究期間：2009 ～ 2011

課題番号：21300096

研究課題名（和文） 検索ログ解析と認知的研究による利用者の探索的な情報検索行動の研究

研究課題名（英文） Users Behaviour during Exploratory Search: Through Log Analysis and Cognitive Studies

研究代表者

神門 典子（KANDO NORIKO）

国立情報学研究所・情報社会相関研究系・教授

研究者番号：80270445

研究成果の概要（和文）：

本研究では、探索的検索の過程における利用者の認知過程と情報探索行動を理解し、モデル化を計るために、発話プロトコル、視線などの詳細な認知的データとログと探索の成果データといった多様な詳細度を持つ利用者の探索過程データを比較分析し、多様な検索の目的やタスク、トピック、利用者の特性に応じた探索過程の特徴を分析した。また、分析のために、Lookzone、Web 行動カテゴリ、LinkDepth、コンセプトマップ、知識変化や知識利用のタクソノミなどの新たな分析枠組みを提案し、データ収集・分析のツールを開発した。合わせて、提案手法を複数の探索的検索システムの評価や設計にも援用し、システム設計の機能や改善の提案を行った。

研究成果の概要（英文）：

We have investigated the users' cognitive process and behavior in exploratory search on the Web. Through the laboratory experiments with more than 200 participants, we have collected users exploratory search process data with different detailness such as think-aloud protocol, gaze, depth interview, questionnaire, search logs, and concept maps, and analyzed users' exploratory search process according to different tasks, topics and users characteristics. We have also proposed a set of new analytical frameworks such as Lookzone, Web action categories, LinkDepth, taxonomies for knowledge changes and utilization, and the data collection and analysis tools such as Copatt, QT honey, viscm maps. The proposed methods were also applied to evaluate various exploratory search systems and system function design.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	6,000,000	1,800,000	7,800,000
2010年度	4,700,000	1,410,000	6,110,000
2011年度	3,200,000	960,000	4,160,000
総計	13,900,000	4,170,000	18,070,000

研究分野：総合領域

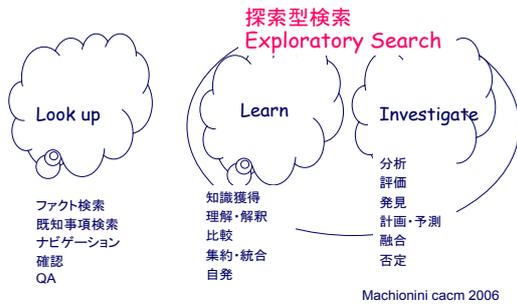
科研費の分科・細目：情報学・図書館情報学・人文社会情報学

キーワード：情報検索・探索行動分析

## 1. 研究開始当初の背景

情報探索は、「参照(Look up)」、「学習(Learn)」、「調査(Investigation)」に分類できるが、既存の Web サーチエンジンは、主として事実検索や既知事項検索などの「参照」を支援し、「学習」と「調査」などの「探索

的検索 (Exploratory Search)」は、近年、急速にその重要性が認識され、関心が高まっているが、その実現メカニズムについての研究は端緒についたばかりである[1]。探索的検索システム構築のためには、利用者の探索過程の理解が必要不可欠である。



情報探索行動研究では、検索の目的や意図、携わっているタスク、背景知識、検索された情報内容などの利用者の「状況(Context)」の重要性が広く認識されているが、少数の利用者を対象として、詳細に情報探索行動やその認知過程を調べた研究が多く、検索システムの設計の提案にはいたらない。

WEB サーチでは、利用者の意図に近づく方策として、クエリログ(Query Logs)における語の共起関係等を用いた検索語想起支援、クリックログ(Click-Through Logs)に基づく検索順位の調整などが、すでに一部で実用化されている。また、クエリログやクリックログなどを用いた「ユーザモデル」研究が急速に発展している。これらのログは、利用者の何らかの行動を表わす、大量に入手可能なデータであり、利用者の行動や状況を検索モデルに組み込むための材料として有望であるが、下記のような問題があり、これだけで利用者の意図や探索過程を理解できるわけではない。

- 1) 検索サーバ側で取得できる検索ログが、利用者の状況や目的とどのように関連しているのかは、ほとんど研究されてなく、実際は何を表わしているのかわからない
- 2) 検索ログの利用は、原則としてサーチエンジン企業の研究者に限られる
- 3) 個人情報保護の問題があり、多くの研究者が共通して利用できないため、システム間の相互比較や有用性の評価ができず、研究が積み重なっていかない。

利用者の探索過程を理解し、それをシステムの設計や評価にも活用できる方策を検討するには、検索ログと利用者の認知過程や探索行動との関連付けを明らかにすることが必要である。

## 2. 研究の目的

本課題では、サーバ側の検索ログ、利用者端末側の操作ログ、詳細な利用者認知実験データなど、詳細さと量が異なる利用者行動記録を相互に関連付けて分析することによって、探索的検索の過程における利用者の認知過程と行動を理解するとともに、タスク、トピック、利用者の属性などが、探索過程に及ぼす影響を検討する。

研究目的は、以下のようにまとめられる。

- 1) 探索的検索の過程における利用者の認知過程と情報探索行動を理解し、モデル

を構築する

- 2) 多様な詳細度と量を持った利用者の情報探索過程データを比較分析し、それぞれの特徴を明らかにする
- 3) 検索ログを探索型検索システム改善に用いることの妥当性・意義・有効な方策を検討する

## 3. 研究の方法

### 3.1 データ収集

探索の目的やタスク、探索内容のトピック、利用者の属性（背景知識、検索やタスクの習熟度、トピックへの親和性など）について、条件を変えながら、以下の方法でデータ収集を行った。

#### 1) 詳細な認知実験

探索的検索を行う実験参加者に対して、探索前後の Depth Interview、視線計測、端末側の操作ログ、画面キャプチャ、発話思考法（探索時、視線つき画面キャプチャによる事後刺激想起、参加者のペアリングなど）、ビデオを用いて、詳細な認知データと行動データを収集した。

#### 2) 端末側ログ収集

端末側ログ収集ツール QT Honey を開発し、実験参加者の探索時の端末側操作ログを収集した。合わせて、米国カーネギーメロン大学で開発し、本プロジェクトで改良に協力した Lemur Querylog Toolbar、研究代表者が関わる他プロジェクトで開発した QuickEdit も援用した。

#### 3) サーチェンジン検索ログ

情報検索の国際評価 CLEF のログ分析研究部門 LogCLEF に参加し、参加研究団体に配布されるサーバ側ログを用いた。

### 3.2 データの分析

収集した各種の認知・行動データを分析するため、下記のような分析枠組みを提案し、分析ツールを開発した。

#### 1) 新規に提案した分析枠組み：

- Lookzone : Web ページ中の見た箇所
- Web 行動カテゴリ : 探索の操作を定義
- Link Depth : リンクをたどる深さ
- コンセプトマップ : 探索の前後の知識構造を比較。教育学での実践を援用し、新規な用法を提案
- 知識変化・知識利用タクソノミ : グラウンデッドセオリーの「絶えざる比較法」によって発話思考法による発話データを分析して得た内容分析枠組み

#### 2) 収集・分析ツールの開発：

- COPATT : 画面キャプチャ、視線、発話データを統合するプラットフォーム
- QT-Honey : クライアント側の操作ログ記録・収集
- Scanpath2SVG : どのランクまで見たか

を可視化する

- CmapAnalyzer, VisCmap: 検索前後のコンセプトマップ対比と可視化
- 探索履歴の可視化ツール: 探索空間上に各利用者の探索・閲覧履歴を可視化

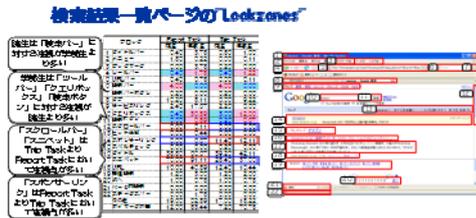


図 1. 検索結果一覧ページ LookZone 分析結果

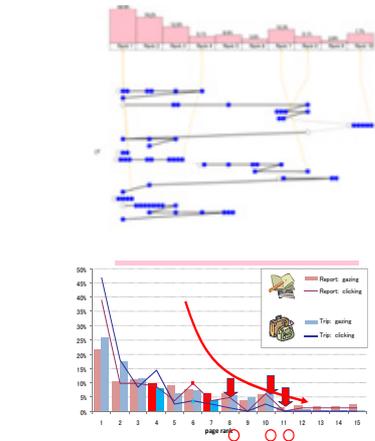


図 2. クリックより下位まで見ている。タスクや利用者の Expertise により傾向が異なる。



図 3. LinkDepth リンクの辿り方は、タスクや利用者の Expertise により異なる

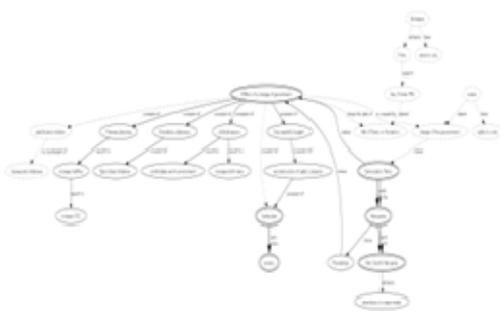


図 4. 探索前後の利用者の知識構造の変化: タスクや Expertise で異なる傾向がみられた

#### 4. 研究成果

利用者の探索過程を記述する手法を提案し、関連ツールを開発し、公開した。また、それらを用い、詳細な認知実験とログと探索の成果(outcome)としての各利用者の知識構造の変化の分析を関連づけ、利用者の特性やタスクに応じて、特徴的な探索過程を明らかにした。また、ログによって、探索過程の特徴を記述する手法を提案し、検証した。

探索過程で、利用者が想定している知識体系がテーマによって切替・詳細化する過程を明らかにした。また、利用者の特性やタスクに応じて、探索すべき情報のサブカテゴリが明確か否かが異なった。それにより、探索戦略やタスクに応じたメタ認知がスムーズな探索のキーであることが示唆された。

探索における対話のモデルとして、コミュニティ QA サイトに着目し、対話を分析し、利用者にとって好ましい解答の特性を明らかにした。

本研究で提案した分析手法を、WEB、多言語ニュース、音楽、画像などを探る各種の探索的検索システムの有効性評価に適用し、検索結果を、クエリに対する出力 (Output) ではなく、探索過程を通じて学んだ成果 (Outcome) としてとらえ、コンセプトマップによって検索前後の知識構造の変化を分析した。個々のクエリ・結果対の評価では、次の探索アクションの多様な選択肢を提示することが柔軟な探索を効果的に支援する観点から、検索結果の多様性 (Diversity) に着目した。さらに、利用者の視点から有効性を評価する指標として、入力クエリ、クリック、リンクナビゲーション、閲覧ページ、閲覧時間 (dwell time)、ブックマーク、トピックやタスクとの親和性などを用いた。

最後に、探索過程における個々の利用者の探索空間を可視化し、個々の利用者の特性や協調的探索の有効性を検証した。

#### 5. 主な発表論文等

(研究代表者及び連携研究者には下線)

【学術雑誌 査読付き論文】(計 5 件)

- ① 大塚 淳史, 関 洋平, 神門 典子, 佐藤 哲司: 「コミュニティ QA を用いたクエリ拡張のためのコンテキスト抽出に関する一考」, DBSJ Journal 11(1) (2012.06) [査読有]
- ② 大塚 淳史, 関 洋平, 神門 典子, 佐藤 哲司: 「情報要求の言語化を支援するクエリ拡張型 Web 検索システムに関する一考察」, 情報処理学会論文誌データベース 51(10) 1234-1244 (2011.10) [査読有]
- ③ M. Miwa, Y. Egusa, H. Saito, M. Takaku, H. Terai, N. Kando, "A method to capture information encountering

embedded in exploratory web searches”. Information Research, Vol.16., no.3, paper487, Sept. 2011. [査読有]

- ④ 加藤 恒昭, 松下 光範, 神門 典子. 「時系列情報の値と変化に関する言語表現コーパスの構築」, 人工知能学会論文誌, Vol. 25, No.5, pp. 637-650, (2010.10) [査読有]
- ⑤ 高久 雅生, 江草 由佳, 寺井 仁, 齋藤 ひとみ, 三輪 眞木子, 神門 典子. 「タスク種別とユーザ特性の違いが Web 情報探索行動に与える影響: 眼球運動データおよび閲覧行動ログを用いた分析」, 情報知識学会誌 vol.20, No.3, pp.249-276, (2010.10) 【情報知識学会論文賞受賞】 [査読有]

#### [学会発表] (計 43 件)

1. X. Hu, N. Kando: “User-Centered Measures vs. System Effectiveness in Finding Similar Songs”, In Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR 2012) Porto, Portugal, Oct. 2012 [査読有]
2. V. Garkavijs, M. Toshima, N. Kando: “GLASE 0.1: Eyes Tell More Than Mice”, Proceedings of the 35th Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval (ACM-SIGIR 2012), Portland, OR, U.S.A., Aug 2012 [査読有]
3. H. Saito, Y. Egusa, M. Takaku, M. Miwa, N. Kando: “Using Concept Map to Evaluate Learning by Searching”, In Proceedings of the 34th Annual Meeting of the Cognitive Science Society (CogSci 2012), 札幌, Aug 2012 [査読有]
4. A. Otsuka, Y. Seki, N. Kando, T. Satoh: “QAque: Faceted Query Expansion Techniques for Exploratory Search using Community QA Resources”, Proceedings of Community Question Answering Workshop at WWW 2012, Lyon, France, Apr 2012 [査読有]
5. V. Garkavijs, M. Toshima, N. Kando: 「提示型ユーザーインターフェースにおける視線計測器を用いた適合フィードバック」, 情報処理学会研究会報告 2011 (IFAT-105) 1-8, 福岡 (2012.03) [査読無]
6. 枝 隼也, 島田 諭, 関 洋平, 神門 典子, 佐藤 哲司: 「話題空間の構成に基づく Web 探索履歴可視化手法の提案」, 第4回データ工学と情報マネジメントに関するフォーラム論文集 C9-1 神戸(2012.03) [査読無]
7. 香川 雄一, 島田 諭, 神門 典子, 佐藤 哲司: 「コミュニティ QA における質問・回答が互いに及ぼす影響の可視化に関する検討」, 第4回データ工学と情報マネジメントに関するフォーラム論文集 C3-3 神戸 (2012.03) [査読無]
8. 大塚 淳史, 関 洋平, 神門 典子, 佐藤 哲司: 「コンテキスト切替による多様な情報要求に対する Web 検索手法の提案」, 第4回データ工学と情報マネジメントに関するフォーラム論文集 F8-4 神戸(2012.03) [査読無]
9. M. Yoshioka, N. Kando, Y. Seki: “Evaluation of Interactive Information Access System using Concept Map”. In Proceedings of the 4th International Workshop on Evaluating Information Access (EVIA), 東京, Dec. 2011. [査読有]
10. D. Ishikawa, N. Kando, T. Sakai: “What Makes a Good Answer in Community Question Answering? An Analysis of Assessors’ Criteria”, In Proceedings of the 4th International Workshop on Evaluating Information Access (EVIA2011) 14-19, 東京 Dec.2011 [査読有]
11. 高久 雅生: 「学習型情報探索行動を探る～CRES 共同研究プロジェクトの成果から～」, 人工知能学会 情報編纂研究会 第7回研究会, 横浜(2011.12) [招待講演]
12. 加藤 恒昭, 松下 光範, 上保 秀夫, 神門 典子: 「NTCIR-9 VisEx の概要: Overview of the VisEx task at NTCIR-9」, 人工知能学会第7回情報編纂研究会, 横浜(2011.12) [査読無]
13. 江草由佳, 齋藤ひとみ, 中島 諒, 高久 雅生, 神門典子, 三輪眞木子: 「Web 情報探索前後のコンセプトマップの分析: 探索者によるノード同定結果を用いて」, 情報処理学会情報アクセス研究会研究報告 2011-IFAT104-4 p.1-8, 福岡 (2011.11) [査読無]
14. X. Hu, N. Kando: “User Evaluation of an Interactive Music Information Retrieval System”, Proceedings of the 5th Human Computer Interaction and Information Retrieval Workshop (HCIR 2011) (poster), Mountain View, CA, U.S.A., Oct. 2011 [査読有]
15. H. Saito, R. Nakashima, Y. Egusa, M. Takaku, H. Terai, M. Miwa, N. Kando: “Changes in users’ knowledge structures before and after web search

- on a topic: analysis using the concept map (Poster)”, Proceedings of the 2011 Annual Meeting of the American Society for Information Science and Technology, New Orleans, LA, U.S.A., Oct. 2011 [査読有]
16. 吉岡 真治, 神門 典子, 関 洋平:「複数国の新聞サイトを比較分析するNSContrastの実験的分析」, 情報処理学会情報アクセス研究会研究報告2011-IFAT103-2 p. 1, 京都 (2011.08) [査読無]
  17. 大塚 淳史, 関 洋平, 神門 典子, 佐藤 哲司:「情報要求の言語化を支援するコンテキスト提示型クエリ拡張法の提案と評価」, マルチメディア, 分散, 協調とモバイルシンポジウム, 京都 (2011.07)
  18. 高久 雅生, 江草 由佳, 寺井 仁, 齋藤 ひとみ, 三輪 真木子, 神門 典子:「タスク種別とユーザ特性の違いがWeb情報探索行動に与える影響: 眼球運動データおよび閲覧行動ログを用いた分析」, 2011 年度情報知識学会年次大会; 高松 (2011.05) [論文賞受賞記念講演]
  19. 齋藤 ひとみ, 江草 由佳, 高久 雅生, 寺井 仁, 三輪 真木子, 神門 典子:「情報検索の前後におけるユーザの知識構造の変化: コンセプトマップを使った分析」, 情報知識学会情報知識学会誌 21(2), 高松 (2011.05) [査読無]
  20. 渡邊直人, 島田諭, 関洋平, 神門典子, 佐藤哲司:「コミュニティ QA における質問の多面的評価法の検討」, 情報知識学会情報知識学会誌 21(2) 163-168, 高松 (2011.05) [査読無]
  21. 石川 大介, 酒井 哲也, 栗山 和子, 関 洋平, 神門 典子:「複数判定者によるコミュニティ QA の良質回答の判定」, 情報知識学会情報知識学会誌 169-177, 高松 (2011.05) [査読無]
  22. 神門 典子:「知的コミュニティ基盤としての検索技術: 文化財デジタルコンテンツ教育目的応用の事例を通して」, 筑波大学 知的コミュニティ基盤研究センター公開シンポジウム, つくば (2011.03) [招待講演]
  23. 大塚 淳史, 関 洋平, 神門 典子, 佐藤 哲司:「情報要求の言語化を支援するクエリ拡張型 Web 検索システム」, 第 3 回データ工学と情報マネジメントに関するフォーラム 論文集 F6-3, 伊豆 (2011.02) [査読無]
  24. 渡邊 直人, 島田 諭, 関 洋平, 神門 典子, 佐藤 哲司:「QA コミュニティにおける質問者の期待に基づく質問分類に関する一検討」, 第 3 回データ工学と情報マネジメントに関するフォーラム 論文集 B5-1, 伊豆 (2011.02)
  25. H. Saito, M. Takaku, Y. Egusa, H. Terai, M. Miwa, N. Kando: “Connecting qualitative and quantitative analysis of Web search process: Analysis using Search Units”, In Proceedings of The 6th Asia Information Retrieval Societies Conference (AIRS 2010), National Taiwan University, Taipei, Taiwan ROC, Dec. 2010, pp. 173-182. [査読有]
  26. 栗山 和子, 神門 典子:「Q&A サイトにおける質問と回答の分析 (4) 質問タイプ分類の一致度について」情報処理学会情報アクセス研究会研究報告 2010-IFAT100-4 p. 1-8, 草津 (2010.11) [査読無]
  27. 神門 典子:「検索ログ解析と認知的研究による利用者の探索的な情報検索行動の研究」, 『言語処理技術の深化と理論・応用の新展開』 2010 年度科研・合同シンポジウム, 東京 (2010.10) [招待講演]
  28. M. Takaku, Y. Egusa, H. Saito, D. Ishikawa, N. Kando\*, H. Terai, M. Miwa: “CRES at LogCLEF 2010: Toward Understanding User Behaviors through Analysis of Search Sessions, Search Units, and Click Ranks ” , In Proceedings of Conference on Multilingual and Multimodal Information Access Evaluation (CLEF 2010), Padua, Italy, Sept. 2010, pp. 1-11 [査読無]
  29. Y. Egusa, H. Saito, M. Takaku, H. Terai, M. Miwa, N. Kando, “Using a Concept Map to Evaluate Exploratory Search” , In Proceedings of the Third International Conference on Information Interaction in Context (IIiX 2010), New Brunswick, NJ, U.S.A., Aug 2010, pp. 175-184. [査読有]
  30. T. Kato, M. Matsushita, N. Kando, “Bridging Evaluations: Inspiration from Dialogue System Research” , In Proceedings of ACM-SIGIR 2010 Workshop on Workshop on Simulation of Interaction: Automated Evaluation of Interactive Information Retrieval, Geneva, Switzerland, July 2010 [査読有]
  31. Y. Egusa, H. Saito, M. Takaku, H. Terai, M. Miwa, N. Kando: “Link Depth: Measuring How Far Searchers Explore Web” , In Proceedings of 43nd Hawaii International Conference on System

- Sciences (HICSS 2010), Koloa, HI, U. S. A., Jan. 2010 [査読有]
32. 栗山和子, 神門典子: 「Q&A サイトにおける質問と回答の分析(3) 利用者の質問・回答履歴を中心に」, 情報処理学会情報学基礎研究会研究報告 2009-FI-97 1-8, 東京(2010.01) [査読無]
33. M. Miwa, S. Erdelez, H. Lund, J. Gwizdka: "What Can Eye-Trackers Visualize?—An Approach to Capture the Reality of Search Processes", Presented at 2009 Annual Meeting of the American Society for Information Science and Technology (ASIS&T), Vancouver, Canada(2009.11) [Panelist]
34. 江草 由佳: 「目は口ほどにものを言う? Web 情報探索行動における視線情報の分析」, 第 6 回 ARG カフェ, 横浜(2009.11) [講演]
35. 三輪眞木子, 江草由佳, 齋藤ひとみ, 高久雅生, 寺井仁, 神門典子: 「Web 上の Exploratory search の特徴: 発話プロトコルと事後インタビュー分析結果より」, 情報処理学会 第 96 回情報処理学会情報学基礎研究会報告会 2009-FI-96 p. 8 東京(2009.11) [査読無]
36. 栗山和子, 神門典子: 「Q&A サイトにおける質問と回答の分析(2)」, 情報処理学会 情報学基礎研究会研究報告 2009-FI-96 1-8, 東京(2009.11) [査読無]
37. H. Saito, H. Terai, Y. Egusa, M. Takaku, M. Miwa, N. Kando: "How Task Types and User Experiences Affect Information Seeking Behavior on the Web: Using Eye-tracking and Client-side Search Logs", In Proceedings of the ACM-SIGIR Workshop on Understanding the User (UUIR 2009), Boston, MA, U. S. A., July 2009 [査読有]
38. 齋藤 ひとみ: 「ユーザの認知プロセスに着目した Web インタラクションの分析」. 第 15 回 Web インテリジェンスとインタラクション研究会会議論文集, 広島(2009.07) [特別講演]
39. 江草由佳, 高久雅生, 齋藤ひとみ, 寺井仁, 三輪眞木子, 神門典子: 「Link Depth: Web 情報探索行動の閲覧パターンの分析」, 情報処理学会 情報学基礎研究会研究報告 2009-FI-95 1-7, 神戸(2009.07) [査読無]
40. 栗山和子, 神門典子: 「Q&A サイトにおける質問と回答の分析」, 情報処理学会 情報学基礎研究会研究報告 2009-FI-95 1-8, 神戸(2009.07) [査読無]
41. N. Kando: "Towards Describing and Capturing the Sentiment in and about Media". CHORUS Workshop on Affect,

- Appeal, and Sentiment as Factors Influencing Interaction with Multimedia Information Access, Brussels, Belgium (2009.05) [講演]
42. 高久雅生, 江草由佳, 寺井仁, 齋藤ひとみ, 三輪眞木子, 神門典子: 「サーチエンジン検索結果ページにおける視線情報の分析」情報知識学会誌 Vol. 19, No. 2, pp. 224-235, 東京(2009.05) [査読無]
43. 齋藤ひとみ: 「発話・行動・視線からみた Web 情報探索行動」, 日本知能情報フレンジイ学会 関西支部例会, 吹田(2009.04)

#### 【図書】(計 2 件)

- ① 三輪眞木子『情報行動: システム志向から利用者志向へ』勉誠出版, 205p. (2012.03)
- ② 三輪眞木子「情報探索行動」, In『情報アクセスの新たな展開: 情報検索・利用の最新動向』. 小山憲司, 岸田和明(編). 勉誠出版. pp. 51-69. (2009.10)

【その他】ホームページ

<http://cres.jpn.org/?FrontPage>

#### 6. 研究組織

##### (1) 研究代表者

神門 典子 (KANDO NORIKO)

国立情報学研究所・情報社会相関研究系・教授

研究者番号: 80270445

##### (3) 連携研究者

三輪 眞木子 (MAKIKO MIWA)

放送大学・ICT 活用・遠隔教育センター・教授

研究者番号: 90333541

齋藤 ひとみ (SAITO HITOMI)

愛知教育大学・教育学部・准教授

研究者番号: 00378233

江草 由佳 (EGUSA YUKA)

国立教育政策研究所・教育研究情報センター・総括研究官

研究者番号: 60413902

高久 雅生 (TAKAKU MASAO)

独立行政法人物質・材料研究機構・科学情報室・主任エンジニア

研究者番号: 00399271

寺井 仁 (TERAI HITOSHI)

名古屋大学大学院・情報科学研究科・特任准教授

研究者番号: 30397442 (H21 のみ)