

令和元年6月24日現在

機関番号：12102

研究種目：基盤研究(B) (一般)

研究期間：2014～2018

課題番号：26280118

研究課題名(和文) 検索メディアと思考パターンに関する検索語生成過程の脳活動データ解析

研究課題名(英文) Brain activity in information search

研究代表者

中山 伸一 (Nakayama, Shin-ichi)

筑波大学・図書館情報メディア系・教授

研究者番号：60155885

交付決定額(研究期間全体)：(直接経費) 12,100,000円

研究成果の概要(和文)：情報検索する際に検索語を考える時の核となる脳内活動ネットワークと、条件によって起動する活動ネットワークが明らかになった。さらには、解析によって、検索対象のメディアに依存しない情報検索の中核に関わる脳機能と、メディアの種類に依存する脳機能を見出すことができた。検索語を考える過程は、それなしでは情報を検索することができず、さらには得られる情報検索結果の質に影響するため、複数ある情報検索行動の過程の中でも、重要性が高い過程である。また、これらの脳活動データからデータベースを構築した。

研究成果の学術的意義や社会的意義

検索語を考える過程は、それなしでは情報を検索することができず、さらには得られる情報検索結果の質に影響するため、複数ある情報検索行動の過程の中でも、重要性が高い過程である。従って、この過程の詳細が明らかになれば、情報検索に関して価値のある知見が得られ、多くの人が日々活用している情報検索サービスの向上に役立つ。本研究の結果によって、空間分解能が高いデータである fMRI 実験データが得られた。

研究成果の概要(英文)：The core brain activity network when generating search query terms was elucidated. It became clear that there is also a network that its activity depends on the search conditions. Furthermore, analyses revealed the core brain functions of information search and the brain functions that depend on the media of information retrieval. The query generation process is important because no information can be retrieved without it. Furthermore, the importance is justified as it influences the quality of the information search results that can be obtained. We also constructed a database of brain activity data.

研究分野：情報学

キーワード：情報検索

## 1. 研究開始当初の背景

(1) Google に代表されるインターネットで利用可能な検索システムは、インターネット上の膨大な情報を検索する方法として、その重要性が増しつつある。これらの情報検索の行動は、何かを得るために探す、という日常生活において重要な行動の1つである。情報検索は、推論や意思決定、記憶や連想、視聴覚等の感覚系からの認知、のように様々な脳の機能が関与し、情報検索の中核活動は脳内で生じている可能性が高いため、従来の研究で行われている外面的な観測のみでは本質を理解することが困難な行動である。そこで本研究では、情報検索に関わる推論や記憶、意思決定といった脳が司る機能を脳活動イメージングによって直接計測し、情報検索の本質を探ることを目的とする。この目的に沿って、申請者らは研究を進めてきた。本研究プロジェクト開始までに、51人の被験者を対象に、情報検索時の脳活動データを計測した。解析の結果、前頭回、眼窩回、尾状核が検索語の生成に関わる主な領域であることが明らかになった。この研究成果は、検索対象のメディアが文章の場合である。しかし、文章以外に画像のように様々なメディアが存在する。

(2) 情報検索行動の重要な過程として、(1) 検索語を考える過程と、(2) 検索結果の適合性を判定する過程が挙げられる。どちらの過程も、被験者の行動を外部から観察する行動実験では多く研究されているが、本研究は、検索語を考える過程に焦点を当てる。その理由は以下の通りである。(i) 検索語を考える過程の方が、検索結果の適合性判定と比較して、従来のように外部観測のみで扱うことが困難であり、脳活動の計測がより重要である。(ii) 現時点で、文字情報の検索だけでなく、画像や動画の検索サービスも提供されているが、キーワードを検索語として入力する場合が殆どである。スマートフォンで提供される音声入力による検索や問合せでも、言葉を入力していることに変わりない。

(3) 現在、脳活動データの計測手法として、fMRI、MEG、EEG、NIRS、PET 等が挙げられ、それぞれに利点と欠点がある。本研究では fMRI を用いて、空間分解能が高いデータを生成し、解析を行う。本研究で対象とする情報検索の作業中には、脳の様々な機能が関与するとともに、使われる機能が時系列で変化すると考えられる。これらの一連の動作の遂行中に、脳活動データを平行して計測することで、どの部位がどのように活動しているかが判る。このデータを、これまでに発表および蓄積されている脳活動データと比較することで、脳活動の観点から情報検索を捉えることができる。

## 2. 研究の目的

(1) 本研究の目的は、情報行動の中核部分について脳活動を計測することである。本研究は、情報検索行動を、脳活動データから記述し、モデルを構築する点、情報検索内容の専門性を要因として変動させて脳活動データを計測し、情報検索の基本的な脳活動パターンを調べる点、である。これらは、これまでは外的側面からしか調べていなかった情報検索行動の研究とは根本的に異なり、情報検索の中核を担うと考えられる脳内の活動から調べるものである。また、本研究の成果として、脳科学の観点からの情報検索モデルが挙げられる。特に、本研究によって得られる脳活動データの時系列データから、従来では切り分けが困難であった検索語の生成過程をより厳密に定義できる。

(2) 本研究の意義としては、外的側面のみを対象とする従来の研究では明らかにできなかった情報検索時の脳活動を捉えること、脳科学という従来とは異なる側面から情報検索および情報行動を捉えること、まとまった数の被験者を対象に、同一条件下で実施する情報検索における脳活動データが得られ、今後の脳研究および情報検索の研究の発展に貢献できること、グーグル等の検索サービスの普及からも判るように、特にインターネット上で増加しつつある情報を有効に活用するための知見が得られること、脳活動イメージングによって情報検索行動を解析する方法論の確立に貢献できること、が挙げられる。

(3) さらに、ランキングとの関連が挙げられる。サーチエンジンの大規模なログ分析から検索結果のクリック頻度は1位のサイトが突出しているが、これが検索されたサイトの関連性によるものであるという意見と、サーチエンジンへの信頼性（半盲目的に一番推薦されたものを選択）という意見がある。脳の活動を見ることで、検索結果の評価とランキングの関係がより明らかになることが期待される。

(4) 最後に感性情報検索との関連性を述べる。生理学やコンピュータビジョンで開発された様々なセンサーを利用して、情報検索行動中のユーザの感情を検証する研究が近年活発になっている。例えば、顔の表情分析から閲覧している文献の主題適合度を推測する研究や、検索の難易度をコントロールすることで感情の変化を調査した例がある。情報検索行動における感性の動きと脳の動きの関係も興味深い研究対象であり、本研究の展開が期待できる。

### 3. 研究の方法

(1) 使用した fMRI 装置は ATR-BAIC に設置されている。fMRI 装置の静磁場強度は 3T、MEG 装置の磁束密度分解能は  $10\text{fT}/\text{Hz}^{1/2}$  以下であり、どちらも本研究の目的には十分な性能を持っている。刺激提示装置としてモニター、そして被験者からの入力装置として、ボタンパッドおよびマイクを用いた。

(2) 得られた脳活動実験データから、検索語の生成時に活動している認知機能を特定した。これまでに蓄積されている膨大な脳研究データを用いた。fMRI データの解析には、MATLAB のツールを用いた。

### 4. 研究成果

(1) 1人の被験者の脳活動計測時間は約40分である。検索トピックが書かれた文章を被験者に提示し、検索語を思考する過程の脳活動を計測する。検索対象とする文章コレクションは、独自に構築した一般的な話題かつ時事性を強く持たない内容で構成される。新たに文章コレクションを構築した理由は、研究開始時に、その内容が本研究に適していると判断して選択した文章コレクションが、時事性が強く、本研究の実験形態と、実験に参加する被験者には向かないことが、予備実験の段階で明らかになったためである。また、検索結果の文章の文字数が数十文字と短かすぎるデータや、表の内容がそのまま文章となっているデータのように、文章の質が一定ではない問題があった。このように、既存の文章コレクションでは、本研究で必要とする最低限の質が確保されておらず、文章コレクションの中から基準以下のデータを削除するには、部分的に人手に頼る必要せいがあるために半自動でしか処理できず、さらには部分的にデータを削除することで、トピック毎の適合文章の割合の変動や実験の精度に影響する可能性がある。このような理由から、本研究で要求される条件を満たす文章を5,000収集し、20個の時事性が弱いトピックを設定した。その後、1つの文章につきトピックの適合性を2人が独立に判断した。このようにして構築した文章コレクションを、脳活動実験で使用した。

(2) 1人の被験者の脳活動計測時間は約40分である。検索トピックが書かれた画像を被験者に提示し、検索語を思考する過程の脳活動を計測する。検索対象とする画像コレクションは、独自に構築した一般的な内容かつ時事性を強く持たない内容で構成される。新たに画像コレクションを構築した理由は、研究開始時に、既存の画像データコレクションが、本研究の実験形態には向かないことが、予備実験の段階で明らかになったためである。このような理由から、本研究で要求される条件を満たす画像を5,000収集し、10個の時事性が弱いトピックを設定した。その後、1つの画像

につきトピックの適合性を2人が独立に判断した。このようにして構築した画像コレクションを、脳活動実験で使用した。

(3) 情報検索で脳活動を計測するには、通常の脳活動実験で使用されるシステムをそのまま利用できないため、本研究の実験に適した実験システムを構築した。このシステムでは、実験時に被験者が口頭で伝える1つまたは複数の検索語を入力すると、検索サーバへ検索語を送信し、サーバから送られる検索結果を受け取ると、検索結果のタイトルリストや検索精度を事前に設定したフォーマットに沿って配置し、画像をリアルタイムで生成するシステムである。また、実験者が操作するインターフェースについては、操作ミスの排除と操作の自動化の向上を、複数のバージョンを経て実現した。この改良によって、被験者が装置内に入っている時間が短縮でき、被験者の負担の軽減と、実験効率の向上が達成できた。このシステムは、今後の情報検索に関わる脳活動研究に有用である。

(4) のべ10人の被験者を対象に予備実験を行ない、検索対象が画像の場合の脳活動実験をデザインした。これは、特に検索語の生成時に重点を置くため、この過程を効果的に計測できる実験である。また、予備実験を実施する過程で、対照実験を複数試し、被験者の応答等から適切な実験内容を決定した。さらには、予備実験によって、被験者への説明手順や実験群と対照群の提示数および提示順序、実験全体の流れ等を、実験データの計測および実施効率の両面から、適切な値を見出せた。本研究でデザインした実験手順は、今後の情報検索における脳活動実験に有用であると考えられる。

(5) 被験者がfMRI装置内で画面に表示される文章を読んでいる状態や、注視点を見ている状態を解析できるようにするため、ビデオカメラを用いてfMRI装置内に入っている被験者の目の状態や眼球運動を録画する。そのために、fMRI装置外に置いたビデオカメラが映像を取得できるように、fMRI装置内に設置する専用の治具を作成した。この治具は、fMRI実験時に眼球運動を記録する必要がある場合に有用である。

(6) 予備実験で確定した実験を実施した。これらのfMRI実験データの解析を進めた結果、情報検索するメディアが文章の場合において、検索語を考える際の核となる脳内活動ネットワークと、条件によって起動する活動ネットワークが明らかになった。検索語を考える過程は、それなしでは情報を検索することができず、さらには得られる情報検索結果の質に影響するため、複数ある情報検索行動の過程の中でも、重要性が高い過程である。従って、この過程の詳細が明らかになれば、情報検索に関して価値のある知見が得られ、多くの人が日々活用している情報検索サービスの向上に役立つ。この解析結果は、空間分解能が高いデータであるfMRI実験データから得たものである。解析によって、検索対象のメディアに依存しない情報検索の中核に関わる脳機能と、メディアの種類に依存する脳機能を見出すことができた。また、これらの脳活動データからデータベースを構築した。

(7) 記述に用いる表現モデルは、複数の観点に基づく記述を統合し、かつ新たな視点によって対象を捉えるための基盤となる。複数の視点は、遺伝子や分子レベルの現象を中心に捉える記述と、身体全体の記述のように、記述対象に直結している視点以外に、モデルに組込まれている内容の組み合わせによって生じる、派生的な視点も存在する。実際に利用時には、後者の派生的な視点の方が有用であり、新規の視点を連鎖的に生成することもある。このような要求を満たすために、記述モデルの改良を行なった。特に、リゾームに類似した構造のモデルへの導入と、モデル全体の時系

列解析の方法を構築することで、当初計画では予定しなかった system of systems として扱うことが可能となり、1つのシステム内の要素の変動がどのように他のシステムへ波及するかの解析方法が見出した。これは、要素単位での変動ではなく、任意の視点によって規定される要素群単位での影響の解析である。これによって、より自由度の高い異なる粒度および抽象レベルの表現と解析が可能となった。

#### 5. 主な発表論文等

[雑誌論文] (計 8 件)

① Hai-Tao Yu, Adam Jatowt, Hideo Joho, Joemon M. Jose, Xiao Yang, and Long Chen. 2019. WassRank: Listwise Document Ranking Using Optimal Transport Theory. Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining (WSDM '19), 査読あり, 24-32.

② Cathal Gurrin, Klaus Schoeffmann, Hideo Joho, Bernd Munzer, Rami Albatal, Frank Hopfgartner, Liting Zhou, Duc-Tien Dang-Nguyen (2019) A Test Collection for Interactive Lifelog Retrieval, MultiMedia Modeling. MMM 2019. Lecture Notes in Computer Science, vol 11295. 312-324. 査読あり

③ T. Maeshiro, “Framework based on relationship to describe non-hierarchical, boundary-less and multi-perspective phenomena”, SICE Journal of Control, Measurement, and System Integration, 11(5), 381–389, 査読あり, 2018

④ T. Maeshiro, Y. Ozawa and M. Maeshiro, “A System Description Model to integrate multiple facets with quantitative relationships among elements”, Human Interface and the Management of Information, 10905, 531–540, 査読あり, 2018

⑤ T. Maeshiro, “Hypernetwork model to describe human-machine system of systems”, Proceedings of APWConCSE 2018, 7, 査読あり, 2018

⑥ T. Maeshiro, Y. Ozawa and M. Maeshiro, “A System Description Model with Fuzzy Boundaries”, Human Interface and the Management of Information: Supporting Learning, Decision-Making and Collaboration, LNCS 10274, 390-402, 査読あり, 2017

⑦ T. Maeshiro and M. Maeshiro, “A System Description Model without Hierarchical Structure”, LNCS 9734, HIMI 2016, 48–59, 査読あり, 2016

⑧ H. Yamaguchi and T. Maeshiro, “A Mechanism to Control Aggressive Comments in Pseudonym Type Computer Mediated Communications”, Proceedings of HCII, 90-100, 査読あり, 2016

[学会発表] (計 2 件)

① Damiano Spina, Jaime Arguello, Hideo Joho, Julia Kiseleva, and Filip Radlinski. 2019. CAIR'18: Second International Workshop on Conversational Approaches to Information Retrieval at SIGIR 2018. SIGIR Forum 52, 2 (January 2019)

② Jaime Arguello, Filip Radlinski, Hideo Joho, Damiano Spina, and Julia Kiseleva. 2018. Second International Workshop on Conversational Approaches to Information Retrieval (CAIR'18): Workshop at SIGIR 2018. 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18), 2018

[図書] (計 0 件)

[産業財産権]

○出願状況（計 0 件）

○取得状況（計 0 件）

〔その他〕 ホームページ等

<http://life.slis.tsukuba.ac.jp/brain/>

## 6. 研究組織

### (1) 研究分担者

研究分担者氏名：真栄城 哲也

ローマ字氏名：Tetsuya Maeshiro

所属研究機関名：筑波大学

部局名：図書館情報メディア系

職名：准教授

研究者番号（8 桁）：30361356

研究分担者氏名：北村 達也

ローマ字氏名：Tatsuya Kitamura

所属研究機関名：甲南大学

部局名：知能情報学部

職名：教授

研究者番号（8 桁）：60293594

研究分担者氏名：上保 秀夫

ローマ字氏名：Hideo Joho

所属研究機関名：筑波大学

部局名：図書館情報メディア系

職名：准教授

研究者番号（8 桁）：00571184

※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、研究者個人に帰属されます。