# 大規模言語モデルに基づいた対話型 AI による研究支援に関 する初歩的分析

田島 逸郎<sup>1,a)</sup>

概要:大規模言語モデル(LLM)は、ChatGPTやBing Chatのような人間との「自然な」対話を生成するように進歩している.LLMは、教育や学術研究などを含め、様々な用途に利用しうる.しかし、LLMがどのようなタスクを実行できるかは明らかではなく、探索的な研究が必要である.本研究では、AIに人間による文書に対して質問や問題点の指摘をさせる大学のゼミのような方法を提示し、またそれがどのように学術研究に貢献できるかを調査することを目指す。まず、AIと人間の実際の対話を収集した。その上で、その対話をエスノメソドロジー的ワークの研究の方針に従って分析し、人間とAIとの対話がどのように学術研究の場となりうるか、特に、(1)応答がいかにして学術的なものとして、さらに(2)学術研究の発展に貢献できるようなものとして受け取られるかを分析した。その結果、対話の中でのAIによる提案には(1)文書を明確にするもの、(2)主題に対するより深い洞察を求めるもの、の2種類あることがわかった。また、AIは人間が書いたものに対して疑問を持たないことがわかった。結論として、AIと文書に基づいて対話することは、文書をより良いものにすることにつながるが、そのためには専門分野についての正しい知識が必要である。

キーワード: 質問生成, 思考・推敲支援技術, エスノメソドロジー

# A preliminary analysis of research support by dialog AIs based on Large Language Models

Itsuro Tajima<sup>1,a)</sup>

Abstract: Large Language Models (LLMs) are improving to generate "natural" dialog with humans, such as ChatGPT or Bing Chat. They can be used for many purposes, including education and academic research. However, what kind of tasks they can perform is not clear and needs to be explored. In this research, I aim to show the seminar-like method that makes AI ask questions or point out problems in the human document, and investigate how it can improve our academic research. I collected a real dialog between AI and a human. And I analyzed a dialogue according to ethnomethodological studies of work to explore how dialogues between a human and an AI can be a scene of academic research. In particular, I explore (1) how we can understand responses as academic and (2) as a contribution to improving academic research. As a preliminary result, I found two types of suggestions in a dialog: (1) to clarify documents, and (2) to introduce deeper insights into topics. And I found that AIs do not question what humans write. In a conclusion, dialog between AIs about documents is related to making academic documents better, but we need correct knowledge about our disciplines.

 ${\it Keywords:}\,$  Question Generation, Writing Support, Ethnomethodology

## 1. 序論

情報処理学会の論文投稿テンプレート [1] に「他の人に

<sup>1</sup> 合同会社 Georepublic Japan

a) niryuu@gmail.com

読んでもらう」という項目があるように、学術研究のあらゆる段階において、他者からのフィードバックは重要である。しかし、多分野にまたがる研究や、新たな分野において知見を発表する段階においては、適切な専門分野の研究者が知人にいるとは限らない。では、対話型 AI にフィードバックをさせた場合、それは研究や研究をまとめる上で役に立つのだろうか。

Bing Chat, ChatGPT などの大規模言語モデル(Large Language Models)に基づく対話型 AI は、2022 年に登場して以降様々な反響を呼んだ。特に、その汎用性から様々なタスクに応用できる可能性があるため、何に活用できるかどうかに関する議論が Twitter などで広く活発になされている [2]. 高等教育や学術研究、ソフトウェア開発などの専門的分野でも例外ではなく、実際に文書生成、質問回答などに用いられているが、それらには情報の信頼性、剽窃などの様々な問題点が指摘されている。また、その効果に関しても相反する分析結果が出ている [3].

一方で、大規模言語モデルはその汎用性から、典型的な「この用途に使うことができる」という枠を超えて活用できる可能性がある.このため、より問題点を克服し、有効に活用するための手法を人間の側が提案する必要がある.また、その手法による AI との対話が実際に人が行っているタスクに対して適切であるかどうかに関する分析も必要である.

よって本研究では、その一つとして人が質問し AI が回答するのではなく、逆に人が学術的な文章を提示しそれに AI が質問や指摘をする手法を提案し、その対話が実際に学術というワークにおいて適切なやり方で行われているかを明らかにすることを目的とする.

#### 関連研究

#### 2.1 大規模言語モデルと対話型 AI

本研究では、対話型 AI を所与のものとし、あくまでその やりとりを人同士の対話とみなして分析する. しかし、対 話型 AI は人とは異なる特徴を持つことも事実であり、それを人は対話の中で理解することができる. このため、対話を分析する上で関連する大規模言語モデルに関する最低 限の前提知識を提示する.

言語モデルとは、与えられた単語の系列から次の系列を出力するモデルである。その中でも、大規模なものが大規模言語モデルであるが、特に本研究で注目する大規模言語モデルは、大量のデータに基づいた大規模なパラメータを持つ、特定のタスクに特化しない一般的な能力を持つ言語モデルを指す[4]. 逆に言えば、どのようなタスクを遂行できるかについて全てを列挙することはできない.

このような大規模言語モデルは、時系列を取り扱う深層 ニューラルネットワークの中でも、前の学習データを必 要としない Transformer モデルの登場によって可能になっ た [5]. すなわち、計算能力さえあれば並列的に学習を行うことができる。一方で、これらのモデルは大規模になるにつれて大量の計算能力を持つ団体でなければ構築が難しくなり、実際にどのようなデータに基づいて、どのようにチューニングされているかは公刊資料だけでは窺い知れない [4].

大規模言語モデルを対話などの特定の用途にチューニングする際には、他の機械学習モデルのようにゼロからチューニングする、もしくは汎用的モデルを特定のタスク及びデータセットに合わせる(ファインチューニング)方法もあるが、その言語モデルが持つ言語理解の機能に基づき、自然言語によって指示を行うことで特定のタスクを行わせる「プロンプトチューニング」でも一定程度の性能を発揮することが可能である [6]. ここで指示に利用する文をプロンプトと呼ぶ. これは対話型 AI にとって重要な発展である. なぜなら、対話型 AI に対するチューニングを対話の中においてすることができるためである.

最後に、大規模言語モデルに基づく対話型 AI の構成について、ChatGPT の原型である InstructGPT[7] を例に挙げて述べる。InstructGPT は、大規模言語モデル GPT-3に対し実際に対話を行い、応答が対話として適切かどうかを強化学習の方法でフィードバックをするという方法で、大規模言語モデルの出力を対話に特化するようにチューニングしている。これにより対話を学んだうえで、プロンプトにより様々な自然言語による指示に対話的に答えることができる。

#### 2.2 質問生成タスク

質問生成(Question Generation)タスクは、与えられた文章に関する質問を生成するタスクである. 質問生成タスクは、答えのあるものとないものに大別される[8]. 例えば事件に関するニュース記事があったとして、事件が起きた日時を問うものが前者で、なぜそのような事件が起こったかを問うものが後者である. 前者は試験問題の自動生成など、後者はニュース記事を補完するもの[8] などの応用例がある.

答えのない (Open-ended) 質問生成は困難なものだとされる [9]. その理由は、答えのある質問が既にある知識の一つの形態であるのに対し、答えのない質問は知らないことを問うことであるのではないかと考えられる. 特に、学術研究における質問は、そもそもプライベートなものであるためデータセットそのものが存在しない. 学術論文に対する質問を集めたデータセットは存在するが、質問はクラウドソーシングにより質問を作るよう指示されたもので、実際に関心を持って質問をしたものではない [10].

以上のように、学術的な質問を行う AI を作ることは、 困難であると考えられ、研究課題として残っている. しか し、大規模言語モデルに基づく対話型 AI を用い、そのよ

うな質問が可能であることを発見した. これが本研究の出 発点である.

# 3. 研究目的

本研究では、学術的文章に対して対話型 AI によるフィードバックをさせることで学術研究を進める手法を提案し、それが実際に学術的なフィードバックであり、また学術にとって役に立ちうるものであるかどうかを解明することを目的とする.

### 4. 研究方法

#### 4.1 データの収集

本研究で分析するデータは、著者自らが本研究の着想を得て学会発表をするに至るまでの過程で主に Bing Chat,補助的に ChatGPT を利用した記録である. 学術的な文章に質問をさせるようなプロンプトを考案し、本研究自体のアブストラクトを随時入力することでその応答を収集した. 主に3つの対話記録を元に分析を行う\*1. なお、個人的な事項も関わるため全てのデータを研究対象とはしていない. このデータの収集方法が適切であるかどうかは、分析手法及び結果において述べる.

#### 4.2 分析手法: エスノメソドロジー的ワークの研究

本研究では、Harold Garfinkel が提唱したエスノメソドロジーの方針 [11] に従い、ある特定の場面、例えば学術研究などで成立している秩序がいかにして成り立っているかについて明らかにする。特に、学術研究のような日常的でない場面においては、その場面特有のやり方で秩序が成立していることがある。その固有性に即した形で秩序の成り立ちを記述するのがワークの研究である。

例えば事例を収集して統計的な分析をする場合,既に多くの事例が秩序だった形で成り立っており,それらが分類可能であることを前提としている.しかし,対話型のAIに関しては,実世界での試行が始まって日が浅く,そもそもAIと「対話」することがいかなることで,いかにして成立していることかは明らかになっていない.それゆえ,ゼミや学会などで実際に起こっている対話と単純に比較することができる段階にもない.このため,AIが出力した文章がいかにして学術的なものであり,さらには学術を建設的に促進させるようなものとして受け取ることができるかについて,その文章や対話の中にある構成から明らかにする.

特に、研究手法としては Anderson と Sharrock が提唱した自らの経験をエスノメソドロジー的に記述する方法である「三人称現象学」[12] に影響を受けている。自らの経験を記述することは、ともすれば単なる感想ともとられがちである。しかし、仮にある作業に従事している人本人でも、

\*1 https://github.com/niryuu/dc2023-data

その秩序の成立を意識しているとは限らず、それを記述するという点で分析の必要性と妥当性が生まれる。その中でAndersonらは、「鍵となる記述(Keyed descriptions)」を重視した。すなわち、ある作業を達成するために行っていることと、達成された秩序の間を埋めるような記述である。このような分析手法の特徴から、本研究は自らの経験をデータとしながらも、個別事例の詳細な分析に重きを置い

このような分析手法の特徴から、本研究は目らの経験を データとしながらも、個別事例の詳細な分析に重きを置い ており、またそれがどれだけの一般性を持つかに関しても 分析の対象となる.

#### 5. 結果

Bing Chat や ChatGPT は試験的サービスの段階にあり (ChatGPT は本稿執筆中に API を有償で提供し始めた), 日々調整が加えられている. それによって, 対話の性質も変わってきた. このため, 結果は主に時系列的に論じる.

#### 5.1 プロンプトによる質問の生成

まず、本研究の探索はそもそも Bing Chat が質問能力を持つかから始まった。 Bing Chat へのアクセスを得た直後の 2023 年 2 月 16 日にいくつかの質問をしたところ、「次にする質問」がサジェストされることを観察した(図 1). これは対話の内容から質問を生成する能力を示しており、2.2 節で述べた質問生成の難しさを既に乗り越えていることも示している。これを受け、質問をするようなプロンプトを考案した。

単に「質問をしてください」では、どのような質問でもすることが可能になる。このため、「専門的」「建設的」「創造的」を制約条件とし、「以下の文章に専門的で建設的で創造的な質問をしてください」というプロンプトを作成した。また、比較のために「以下の文章に専門的で建設的で創造的な指摘をしてください」と文章への指摘を指示するプロンプトも用いた。

#### 5.2 質問の生成と対話

#### 5.2.1 内容に関する質問

同 2023 年 2 月 16 日に、Bing Chat に実際にプロンプトと初期のアブストラクトを入力した。結果を図 2 に示す。生成された 3 つの質問はいずれもトピックが異なるものであるが、「チューニングの要件と方法」「分析手法の立場」「仮設の検証と一般性」といったように、どれも学術研究が備える要件に関連していることを指摘している。またトピックがいずれも与えた学術的な文章に関連しているという点で、生成された質問は与えられた文章に対する「学術的な指摘」ととることができる。一方で、例えば研究者同士の対話ではそれぞれの研究者の専門性や関心に即した形で対話が行われるが、Bing Chat による質問は幅広いトピックに渡っており、特定の方向性を持っていない。どちらかというと学術研究としての適切性に関する質問だと理



図 1 Bing Chat による質問の推薦

Fig. 1 Suggestion of questions in using Bing Chat

以下の文章に専門的で建設的で創造的な質問をしてください。 学術研究におけるAIの活用は、論文の 検索、質問への回答などの形で行われてきた。これらはデータからの学習に基づくものであり、新たな 知見を得るための創造的な活動を支援することは困難であると考えられてきた。しかし、現在登場した 大規模言語モデルに基づくAIは、チュ-ングによって文章に対して建設的で創造的な質問や指摘をあ る程度行うことができる。これは、大学のゼミなどで行われている活動であり、質問や指摘は学術研究 に不可欠である。このため、大学などに所属しない個人研究者にとってAIは有用であるという仮説を立 てることができる。この仮説を検証するため、実際に論文を書いてその過程をエスノメソドロジー的ワ 一クの研究として記述することを試みようとしている。 この文章は、大規模言語モデルに基づくAIが学術研究における創造的な活動を支援する可能性について 述べたものですね。専門的で建設的で創造的な質問をするとしたら、以下のようなものが考えられま • 大規模言語モデルに基づくAIが文章に対して建設的で創造的な質問や指摘を行うためには、どのよ うなチューニングが必要ですか? そのチューニングはどのように行いますか? • 論文を書く過程をエスノメソドロジー的ワークの研究として記述することで、どのような知見が得 られると期待していますか? その知見は学術研究におけるAIの活用方法や効果に関するものですか? • 個人研究者にとってAIは有用であるという仮説を検証するためには、どのような評価指標や方法が 必要ですか? その評価指標や方法は他の研究者と共有できますか?

図 2 Bing Chat によるアブストラクトへの質問

appeared first in https://twitter.com/niryuu/status/1626148308648611842/

Fig. 2 Questions for given abstract from Bing Chat

#### 解できる.

そもそもこれは「質問をする」という行為なのだろうか. 「質問をするとしたら、以下のようなものが考えられます」 と述べているように、Bing Chat の関心に基づいて質問を するというよりむしろ一般に想定される質問を提示してい るととることができる. 本研究では Bing Chat による質問 の提示をそのように捉える.

#### 5.2.2 文章構成に関する質問

米国時間 2023 年 2 月 15 日に、Microsoft 社は Bing Chat の最初の 1 週間の試行についてレポートをまとめた [13]. その結果を受けた改善か、2 月 20 日には内容ではなく文章 構成に関する質問を生成するようになった.

プロンプトと変更を加えたアブストラクトを入力したところ、まず「以下の質問は、文章の内容や目的に関連しています.」と、研究目的や方法、評価基準に関する質問がなされた。全てに答えることはできないため、「エスノメソドロジー的ワークとは何ですか」という質問について回答した。その結果、「エスノメソドロジー的ワークの研究について詳しく教えていただきました.」と返した上で、「AIによる質問がコミュニケーションとして成り立っているかどうかを分析することは、興味深い課題ですね.」と回答をパラフレーズした上で、「文章の構成や論理展開に関連し

ています.」と、論文の議論の流れに関する質問がなされた. その後、「文章の目的や意義に関連」した質問がなされた. 論文が解決する問題や課題、提供する価値や貢献についてである. その上で、データの入手方法、使用した自然言語処理モデル、分析の基準とその妥当性などに関する議論を行うことができた.

このように、2月16日のやりとりと異なり、まず文章構成について順番を追って系統的な質問がなされた。その上で、内容に関する込み入った質問がなされた。ここで、「論文の文章構成に関する質問」と「内容に関する質問」の2種類がなされることがわかった。またこの段階ではまず文章構成が優先されている。

#### 5.2.3 特定の主題に対する掘り下げ

前節の段階では、著者はまだ公的に発表できる段階ではないと考えていた。研究として成立させるため、どちらかといえば「内容」を掘り下げることを求めていた。このため、「指摘」および、ChatGPTによる質問の生成を試行した。同じ内容のアブストラクトに対し「以下の文章に専門的で建設的で創造的な指摘をしてください」と指摘を求めるプロンプトを与えた。その結果、文章構成や、先行研究の問題に関する指摘をされ、それを受け修正を行い「修正しました」と再提出した。

修正において、「プロンプト」に関する説明を付け加えたが、それに関する質問が来た。まず、「これはどのような方法なのでしょうか。」と語義を問う質問が来た。このためプロンプトチューニングの方法について説明したところ、どのようにプロンプトを作成したのか質問が来たため、作成方法について説明した。さらに、そのプロンプトによる出力が実際に有用なものだったか、実例を用いて説明するよう指示されたため、そのように回答した。その後、他の方法との比較や限界について質問され、答えたが、米国時間2月17日に設定されたやり取りの限界に達し、対話を続けることができなくなった。

以上の一連のやりとりにより、研究が登場した背景にとってプロンプトによる手法が可能になったことが重要であること、また実際の対話を分析することの重要性についてあらためて明晰にすることができた。つまり、Bing Chatとの内容面の対話によって研究の学術的な内容全般について深めることができた。その点でプロンプトに関する対話は他の話題に比べても学術を進めるのに有意義なものであった。しかし、そもそもなぜプロンプトに関心を持ったかについてはわからなかった。

#### 5.2.4 Bing Chat の限界や、ChatGPT との比較

前節までで、Bing Chat における主要な質問行為の特徴について述べた。その上で、対話型 AI 特有の特徴に関して、対話の中における Bing Chat の限界と、異なる対話型 AI である ChatGPT との比較について述べる。

まず、Bing Chat と ChatGPT が採用していると考えられる素朴な Transformer モデルの限界として、長いトークンを読み込むと推論速度が急激に下がるという問題がある。このため、一度に入力する文字数の制限や、覚えていられる対話の総量の制限が設定されている、例えば論文全文をもとに議論することなどは難しく、アブストラクトを提示するにとどまった。これに関しては、Longformer[14] などの解決策が提案されている。

次に、Bing Chat と ChatGPT の違いについて述べる. 質問の特徴として、ChatGPT は本研究における「内容に 関する質問」をする傾向にあることがわかった。例えば、 対話型 AI の学術研究に対する有用性、質問生成の難しさ の克服の方法、適切な質問の検証方法などである.

また、対話上の特徴の違いとして、Bing Chat は質問をするよう指示されたら「質問をし、回答をする」ことに固執していたのに対し、ChatGPT はある程度整理されてきたら今までの議論の要約をしたりと、途中で質問を止めてしまった。質問と回答をペアとして見ることは会話分析で隣接ペア [15] と呼ばれる現象である。自然な会話においては質問に対しては回答が来ることが予期され、なかった場合回答の不在が理解できる。それゆえ、Bing Chat が回答に対して次の質問をしてきたということは、相手が前に提示した質問に適切に答えたと、少なくとも対話相手や分析

者にとっては理解できるような形で対話を行っている. 質問と回答という形式を維持することは、考えや文章を深めることにとって重要であるため、本研究の用法では Bing Chat の方が適切であると考えた.

一方で、Bing Chat には同じトピックの質問に固執する側面も見られた.例えば、研究目的について修正を提案した後に修正してもまた同じ修正を提案するなどである.これは前述のトークンの量の限界によるものかもしれない.また、特に指摘するところがなくなったら同じような指摘を何度もすることがあった.これに関しては、「内容面での指摘はありますか」などと別の質問を求めたら、トピックを変えることができた.

#### 5.3 学術的ワークとしての Bing Chat との対話

以上のように、論文のアブストラクトと適切な指示を与えれば、Bing Chat はその内容に即した改善案を提示していた。その点で Bing Chat の質問や指摘は学術的である。また、論文の形式、および内容面での指摘の中には著者が明確に文章の形にしていなかったものもあり、その点でさらなる調査や研究を促すような、研究を進めるものであったと言える。これらから、まず対話型 AI からのフィードバックは研究を進めるというワークに対して適切なものだったと言える。

一方で、これらのフィードバックは研究者同士のインフォーマルな会話、ゼミ報告、学会発表、論文の査読などとは明らかに異なるものであり、特別に高い専門性を有していない、例えば学会発表における質疑応答では、論文の形式面での指摘は行わない、どちらかといえば、対話型 AIによる指摘は「論文の書き方」に関する書籍などに書かれている注意点に近い、しかし、書籍に書かれている利用法は自分で書いた内容を自分で読んでチェックしなければならないのに対し、対話型 AI からはある意味で初めて読む立場からの指摘を受けることができる。その点で、論文執筆を支援できる技術としては有用であり、それゆえに学術研究に資するものである。

最後に、対話型 AI からの指摘の限界の一つとして、「相手が述べたことを否定しない」ということが挙げられる。例えば、プロンプトに対して質問を行った際も、調べてさまざまな用例を調べた上で該当するものがなかったにもかかわらず、「あなたのプロンプトの概念は間違っている」などとは返さなかった。逆に言えば、内容については研究者自らがきちんと学習、研究していることを前提としており、それが本手法が学術研究に有効に活用できる条件だと考えられる。

#### 6. 議論と結論

本研究では、対話型 AI の学術研究への活用可能性の一つとして、学術的文章、特にアブストラクトについて質問

の形式でフィードバックを得ることを提案し、またそれがどのように学術研究に貢献できるかを実際の対話から調査した.その結果、論文執筆の面で、文章をより良くできる手段として活用できること、また活用のためには既に十分に正確な知識を持っていることが必要であることがわかった.宮下らがGPT-3を用いて提案した「不採択コメント」の生成[16]による手法と別の形で学術研究の支援を行うことができるといえ、対話型 AI が潜在的に様々な学術研究の段階において活用できる可能性を示唆している.

本研究の動機の一つとなったものが、奥出直人によるパーソナルコンピュータ初期の論文執筆に関する文献 [17]である. 文献目録、アウトラインプロセッサなど様々な手法が 1980 年代当時から使われ、現在も使われている. しかし、エキスパートシステムに基づく対話型 AI による思考支援システムが紹介されていたが、それは実用的な道具としては途絶えてしまった. では、近年現れた対話型 AI は学術研究に対して何ができるのだろうか. 本研究は、その問いに対して一定の答えを与えた.

#### 6.1 本研究の展望と限界

本研究は,「アブストラクトに対する指摘」という比較的研究の初期段階にのみ焦点を当てた.アブストラクトは研究全体のアウトラインであり,また論理構造の一貫性を示すものでもある.その点でアブストラクトに適切な内容面,構成面の指摘をさせることができたことは,学術研究に対する一般的な認識を対話型 AI が学習していることを示している.しかし,当然学術研究には構想から分析,執筆など様々な段階があり,それぞれが異なる作業である.それらに合わせた支援のあり方がありうる.

また、本研究ではつまり学術研究というワークの中で対話型 AI がいかにして自然に活用できるかに関する基礎を分析した。本研究の結果を受けた他の研究の方向性も可能だろう。例えば被験者を募って提案手法を使って論文を書いてもらうこともありうるが。また、例えば arXiv のデータセット [18] などを利用して様々なプロンプト、アブストラクトに対して質問を生成し、特徴を統計的に分析することも可能だろう。ChatGPT が API を公開したため、全自動でデータ収集を行うことも可能になった。本研究は、それらの今後可能な研究に対して必要な基礎的分析を与えた。

#### 参考文献

- [1] 情報処理学会:情報処理学会研究報告の準備方法 (2018). (論文著者に配布).
- [2] Taecharungroj, V.: "What Can ChatGPT Do?" Analyzing Early Reactions to the Innovative AI Chatbot on Twitter, Big Data and Cognitive Computing, Vol. 7, No. 1, p. 35 (2023).
- [3] Bašić, Ž., Banovac, A., Kružić, I. and Jerković, I.: Better by You, better than Me? ChatGPT-3 as writing assistance in students' essays. (2023).

- [4] Carlini, N., Tramer, F., Wallace, E., Jagielski, M., Herbert-Voss, A., Lee, K., Roberts, A., Brown, T. B., Song, D., Erlingsson, U. et al.: Extracting Training Data from Large Language Models., USENIX Security Symposium, Vol. 6 (2021).
- [5] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. and Polosukhin, I.: Attention is all you need, Advances in neural information processing systems, Vol. 30 (2017).
- [6] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A. et al.: Language models are few-shot learners, Advances in neural information processing systems, Vol. 33, pp. 1877–1901 (2020).
- [7] Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Gray, A. et al.: Training language models to follow instructions with human feedback, Advances in Neural Information Processing Systems (2022).
- [8] Chakrabarty, T., Lewis, J. and Muresan, S.: CONSIS-TENT: Open-Ended Question Generation From News Articles, arXiv preprint arXiv:2210.11536 (2022).
- [9] Kumar, V. et al.: ClarQ: A large-scale and diverse dataset for clarification question generation, arXiv preprint arXiv:2006.05986 (2020).
- [10] Dasigi, P., Lo, K., Beltagy, I., Cohan, A., Smith, N. A. and Gardner, M.: A Dataset of Information-Seeking Questions and Answers Anchored in Research Papers, Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 4599–4610 (2021).
- [11] Garfinkel, H.: Ethnomethodology's Program: Working Out Durkheim's Aphorism, MD: Rowman and Littlefield, Lanham (2002).
- [12] Anderson, R. J. and Sharrock, W. W.: Action at a Distance: Studies in the Practicalities of Executive Management, Routledge, London (2018).
- [13] Microsoft: The new Bing and Edge Learning from our first week, Microsoft (online), available from (https://blogs.bing.com/search/february-2023/Thenew-Bing-Edge-%E2%80%93-Learning-from-our-first-week) (accessed 2023-02-15).
- [14] Guo, M., Ainslie, J., Uthus, D., Ontanon, S., Ni, J., Sung, Y.-H. and Yang, Y.: Longt5: Efficient textto-text transformer for long sequences, arXiv preprint arXiv:2112.07916 (2021).
- [15] Schegloff, E. A. and Sacks, H.: Opening up closings, Semiotica, Vol. 8, pp. 289–327 (1973).
- [16] 宮下芳明,大友千宙: AI が生成する不採択コメントを用いた研究指導,第30回インタラクティブシステムとソフトウェアに関するワークショップ,日本ソフトウェア科学会(2022).
- [17] 奥出直人:物書きがコンピュータに出会うとき,河出書 房新社 (1990).
- [18] Clement, C. B., Bierbaum, M., O'Keeffe, K. P. and Alemi, A. A.: On the Use of ArXiv as a Dataset, arXiv preprint arXiv:1905.00075 (2019).

ここに掲載した著作物の利用に関する注意 本著作物の著作権は情報処理学会に帰属します。本著作物は著作権者である情報 処理学会の許<mark>弱のあったに掲載するものでです。paご</mark>利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」6に従うこ とをお願いいたします。