

eヘルスリテラシーと対話型AIを活用した健康情報の真偽判断との関連

Relationship between e-Health literacy and the Process of Evaluating the Authenticity of Health Information using Interactive AI

東 るみ子

Rumiko AZUMA

日本大学商学部

College of Commerce, Nihon University

Email: azuma.rumiko@nihon-u.ac.jp

あらまし：本稿では、大学生のeヘルスリテラシーと対話型AIを使った健康情報の検索行動および情報の真偽判断との関連について調査を行い、eヘルスリテラシーの高い群と低い群でどのような違いがみられるのかを比較分析した。その結果、質問回数や文字数に違いはみられなかったものの、eヘルスリテラシーが高い学生ほど様々な条件を付与した質問文を作成する傾向にあり、様々な視点による回答を基に真偽の判断を行っていることがわかった。

キーワード：eヘルスリテラシー、対話型AI、ChatGPT、情報リテラシー

1. はじめに

インターネットを利用し膨大な情報の中から信頼できる健康・医療情報を選び出して適切に活用できるスキルをeヘルスリテラシー⁽¹⁾と呼ぶ。eヘルスリテラシーが低いと不正確な情報に基づき、間違えた治療法を選んでしまったり、実際には効果のない不適切な健康行動をとってしまったりするリスクがあるため、情報リテラシー教育と併せてeヘルスリテラシー教育も重要になってくる。本稿の先行研究でも、大学生のeヘルスリテラシーの度合いによって実際に検索エンジンを使った情報検索行動に違いがみられることが明らかとなっている。

一方、2022年11月に米国OpenAI社よりChatGPTがリリースされたことにより、世界的に生成AIの利用が一気に広がり、学生の間でも日頃の学修や課題の取り組みで活用されるようになった。今後は、健康・医療情報の取得においても、検索エンジンを利用した情報検索に代わり生成AIが活用されていくことも予想される。そこで本稿では、大学生のeヘルスリテラシーと対話型AIを使った健康情報の検索行動との関連について調査を行う。先行研究でみられたeヘルスリテラシーの違いによる情報検索行動の違いは、対話型AIを使った場合も同様に検索行動および真偽判断に違いがみられるのかを分析、考察する。

2. 大学生の健康・医療情報の入手方法に関する実態調査

大学生が健康・医療情報をどのような方法で入手しているのか、また対話型AIは活用しているのかの実態を探るため、本学の学生を対象に2024年4月上旬にアンケート調査を実施した。有効回答数は85名（男性67名、女性18名）であった。

学生が日頃から健康情報をどのような手段で調べているのかの質問では、最も多かった回答が「検索エンジンで調べる」（94.1%）であり、次いで「家族

に相談する」（65.9%）であった（表1）。一方、「ChatGPTなどの対話型AIで調べる」の回答は3.5%と低かった。

また、「ChatGPTなどの対話型AIで医療や健康情報を調べたことはありますか？」の質問においても、「はい」と答えた学生の割合は12.9%（11名）と低かった。「いいえ」と回答した学生に理由を尋ねたところ、「Googleなどの検索エンジンで十分だから」や「検索エンジンを使うよりも手間だから」といった検索エンジンが慣れているという理由と、「正しい答えを教えてくれるか不安があるから」や「嘘の情報が混じっているかもしれないから」といったハルシネーションを危惧する理由が大半であった。一方、大学のレポートや課題に対話型AIを使用しているかの質問に対しては、56%の学生が使用していると答えていることから、医療・健康情報においてはより慎重に情報を調べていることがうかがえる結果となった。

表1 「日頃自分の健康状態に問題があるとき、どのような手段で調べますか？」の集計結果（n=85）

選択肢（複数選択可）	割合(%)
検索エンジンで検索し調べる	94.1
家族に相談をする	65.9
YouTubeで調べる	22.4
友達に相談する	20.0
調べずにすぐ病院に行く	12.9
SNSで調べる（相談する）	10.6
ChatGPTなどの対話型AIで調べる	3.5
本で調べる	1.2

3. eヘルスリテラシーと対話型AIを使った健康情報の検索行動・真偽判断の関係

3.1 調査概要

次に、大学生のeヘルスリテラシーの度合いに応じて、対話型AIを使った検索行動および真偽判断に違いがみられるのかを調査をした。調査対象者は、大学2年次以上の20名の学生であった。

調査方法は、学生にこちらが提示した健康問題について ChatGPT (GPT-3.5) を使って実際に調べてもらい、その際に入力した文章と真偽の判断結果について Google Form から回答してもらった。また同時に、光武⁽²⁾によって開発された個々人の e ヘルスリテラシーを数値化し評価できる尺度 (eHEALS 日本語版) と、ChatGPT に関する簡単なアンケートにも回答してもらった。

3.2 e ヘルスリテラシーのアンケート結果と eHEALS の得点について

「医療・健康情報を得る際に利用している情報源」のアンケート結果を表 2 に示す。最も多かったのが「Web サイト」からであり次いで「家族・友人」からであった。これは実際に健康状態に問題がある時に調べる手段を尋ねたアンケート結果 (表 1) と同様の結果であった。また SNS で情報を得ている学生も過半数以上おり、大学生の医療・健康情報の収集手段としてはインターネットが一般的であることがわかった。

表 2 「医療・健康情報を得る際に利用している情報源」の集計結果 (n=20)

選択肢 (複数選択可)	割合 (%)
Web サイト	90.0
家族や友人	85.0
SNS (X や Instagram など)	60.0
テレビ・ラジオ	50.0
新聞	5.0
雑誌	5.0
書籍	5.0
行政からの情報	5.0

調査対象者の eHEALS の平均点は 22.9(SD=7.08) であり、大学生を対象にした先行研究の値(M=22.7, SD=5.1), 及び一般市民を対象とした光武⁽²⁾の研究結果の値(M=23.5, SD=6.5)とほぼ同様であった。本稿では、eHEALS 得点の中央値(Me=22)で高群と低群に分け (中央値は低い群に含め)、対話型 AI を使った健康情報の真偽判断問題について二群による比較を行った。

3.3 対話型 AI を使った健康情報の検索行動および真偽判断の結果

先行研究で実施した「ご飯やパンなどの炭水化物を抜くダイエット方法」の効果に対して、ChatGPT を使って納得のいく結果が出るまで調べてもらい、その結果に対する真偽を判断してもらった。そしてその結果について、eHEALS 高低群で比較を行った (表 3)。eHEALS 高群の学生は「効果なし」と判断した割合が 90.0%だったのに対し、低群の学生では 50.0%と半々であった。先行研究の検索エンジンを使った場合においても本稿の結果と同様に、高群が低群よりも「効果なし」と回答する傾向にあった。しかしながら、今回は高低群間において統計的な有意差はみられなかった。

表 3 健康情報の真偽判断の結果 (n=20)

効果	eHEALS 低群	eHEALS 高群	全体
あり	50.0%	10.0%	30.0%
なし	50.0%	90.0%	70.0%

Fisher's exact test ($p=0.14$)

3.4 真偽判断のためのプロンプト内容の分析

次に、この問題の真偽判断においてどのようなプロンプトを作成し判断したのかを分析した。

質問回数と文字数平均については、eHEALS 高低群における有意差はみられなかった。プロンプトに入力された文に対しては KHcoder⁽³⁾を使いテキストマイニングを行った。表 4 は、各群と関連の強かった特徴語を示したものである。両群とも、1 回目の質問は「炭水化物を抜くダイエットについて教えて」のような問題文に沿った文章であった。しかし 2 回目以降において高群は、特徴語にもみられるように、メリットやデメリットを質問したり、「どういう人が効果が出ないのか」、「子供・高齢者は有効なのか」、「運動と比べて効果があるのか」など、様々なケースや条件を付与した質問内容を送信していた。一方、低群は「炭水化物を摂取しないダイエットを継続して続けるために重要なこと」や「適したダイエットの方法」といったダイエットを成功させるための方法を質問している傾向にあった。

表 4 プロンプト内容の特徴語 (n=20)

	eHEALS 低群		eHEALS 高群
炭水化物	0.366	ダイエット	0.442
ご飯	0.212	効果	0.343
パン	0.206	方法	0.333
重要	0.179	抜く	0.278
人	0.156	デメリット	0.071
摂取	0.143	メリット	0.071
一番	0.107	高齢	0.071
継続	0.107	子供	0.071
適す	0.107	運動	0.069

数値は関連の強さをあらわす Jaccard 係数

4. おわりに

本稿では、大学生の e ヘルスリテラシーと対話型 AI を使った健康情報の検索行動および真偽判断との関連について調査、分析を行った。今後の課題として、データ数を増やし再度分析を行うことが挙げられる。また対話型 AI はプロンプトエンジニアリングの違いにより質および関連性の高い出力が得られるか否かが決まるため、その観点からの比較・分析も課題である。

参考文献

- (1) Norman C.D., Skinner H.: eHealth Literacy: Essential Skills for Consumer Health in a Networked World, Journal of Medical Internet Research, 16;8(2):e9 (2006)
- (2) 光武誠吾, 他: “eHealth Literacy Scale (eHEALS) 日本語版の開発”, 日本公衆衛生雑誌, 58(5), 361-371 (2011)
- (3) 樋口耕一: “社会調査のための計量テキスト分析”, ナカニシヤ出版, (2014)