Vol.2024-HCI-210 No.16
Vol.2024-UBI-84 No.16
2024/11/18

アクセシビリティのための論文内画像の代替テキストに関する定量的調査

宮崎 翔[1,a)]　マラクリヤ シルバン[2,1,b)]　矢谷 浩司[1,c)]

概要：学術論文が視覚に不自由があるユーザにとっても読むことができることは非常に重要であり，中でも，支援技術であるスクリーンリーダーによる読み上げのために，論文内の画像に適切な代替テキストを設定することは非常に大切であるとされている．しかし，現在公開されている論文では，図に適切な代替テキストが設定されていなかったり，設定されていてもその内容が適切でなかったりする場合が少なくない．これに対し本研究では，ヒューマンコンピュータインタラクション分野の国際会議である CHI で発表された 3,000 件以上の論文から，約 17,000 件の論文内の図と代替テキストの組を収集し，その内容について分析を行う．この分析により，現状の代替テキストの設定状況や記述傾向を調査するとともに，今後，論文内画像の代替テキストの記述支援を行う研究につながる知見の提供を目指す．

A Quantitative Study on Alternative Texts for Images in Academic Papers for Accessibility

1. はじめに

学術論文を視覚に不自由のある方でも読めるようにすることは非常に重要であり，そのための取り組みの一つとして，論文内の画像に対して代替テキストを設定することが必要とされている [1]．なぜならば，視覚に不自由のある方はスクリーンリーダーという支援技術を用いて，デジタルデバイス上のコンテンツを聴覚で理解することが一般的であるが，視覚情報である論文内の画像を，聴覚情報としてスクリーンリーダーユーザに伝えるためには，適切な代替テキストが設定されている必要があるからである．

近年，ヒューマンコンピュータインタラクション (HCI) を始めとした研究コミュニティでは，論文のアクセシビリティ向上に関する取り組みの一環として，代替テキストの設定の重要性が高まっており [2]，適切な内容の代替テキストを記述するためのガイドラインの提案などが行われている [3], [4].

しかし，これまでに行われた調査によると，現在公開されている HCI 研究論文においては，図に対する代替テキストが設定されていなかったり，設定されていても短くて不十分であるものが多かったりするなど，依然アクセシブルでないことが指摘されている [1], [5]．そして，代替テキストが適切に設定されていない理由としては，

- 代替テキストを設定する必要性が認識されていないこと
- 代替テキストを記述することが面倒であること
- 代替テキストを記述することが難しいこと

などが考えられる．特に，論文内画像に対する代替テキストの記述の難しさに関しては，その内容の複雑性や文脈依存性の高さから，他の領域における代替テキスト記述と比較して，その特異性と記述の難しさが指摘されている [1], [5].

そして，これらの課題の解決策として，昨今，生成 AI などを活用した代替テキスト記述支援システムの開発が進められている [6], [7] が，その設計において必要とされる知見について，学術論文内の画像に対する代替テキストに関する調査はまだ十分に提供されていない [8].

これに対し本研究では，ヒューマンコンピュータインタラクション分野の国際会議である CHI において，2021 年から 2024 年までに発表されたすべての論文に含まれるす

[1] 東京大学 Interactive Intelligent Systems Laboratory
[2] リール大学 Inria センター
[a)] kakeru-miyazaki@iis-lab.org
[b)] sylvain.malacria@inria.fr
[c)] koji@iis-lab.org

べての画像を収集し，論文内の画像に対する代替テキストの設定状況や記述における傾向を調査する．この研究では，約 17,000 件という，先行研究と比較してかなり大規模なデータに対し，既存のガイドライン [3] に則した定量的な分析を行うことで，現在，著者が代替テキストを記述する際の傾向や注意すべき点を明らかにするとともに，今後の代替テキスト記述支援システムの開発に活用可能な知見を提供する．また，将来的には，今後のアクセシビリティに関する研究に利用可能な大規模なデータセットを提供することを目指す．

2. 関連研究

2.1 代替テキスト記述の重要性の高まり

中等度および重度の視力障害を持つ人は，世界に 2 億 9500 万人いると推定されており [9]，ウェブサイトやアプリケーションをはじめとしたデジタルコンテンツのアクセシビリティは非常に重要である．支援技術の一つであるスクリーンリーダーは，視覚障害のある方がデジタルデバイスを利用する際の支援技術として広く利用されているが，画像に対する代替テキストは，スクリーンリーダーが画像の内容を視覚ではなく聴覚を通して，利用者に伝えるために利用される．そのため，デジタルコンテンツのアクセシビリティ向上のためには，適切な代替テキストを設定することが重要であるとされている [10].

これに対し，学術論文においても，そのアクセシビリティを向上させる取り組みの一環として，代替テキスト記述の重要性が認識されるようになってきており [1]，論文内画像の代替テキストを記述するためのガイドラインの提案などが行われたり [3], [4], [11]，CHI への論文の提出時に代替テキスト記述が実質的に義務づけられる [12] などしている．

なお，本研究で，分析に用いる ACM のガイドライン [3] においては，代替テキストを記述する際に気をつけるべき点として，7 項目を提示しており，それぞれの内容を要約すると以下のようになる．

(1) 代替テキストとしてキャプションを再利用しないこと．
(2) 画像をできるだけ具体的に記述すること．
(3) 代替テキストを短く保つこと．125 文字以下が望ましい．
(4) キーワードを使用すること．
(5) キーワードの過剰使用を避けること．
(6) 代替テキストに「image of」や「picture of」などを含めないこと．
(7) longdesc= ””を忘れないこと．

2.2 代替テキスト記述を行う上での課題

しかし，代替テキストを記述することは，視覚障害のない人にとっては副次的な作業として捉えられがちであり [1]，論文内の画像に対する代替テキストが設定されていない場合が多いことが指摘されている [5]．そして仮に代替テキストを設定したとしても，その内容をアクセシビリティの観点から適切なものにすることは容易ではない [13].

代替テキスト記述の難しさを表す例として，Bigham らの研究 [1] が挙げられる．この研究では，ユーザ実験の一つとして，UIST2015 に論文を提出した著者に対し，報酬の対価として代替テキストを著者に設定するように促す実験を行った結果，全体の 60%以上の代替テキストが，キャプションをコピーしたものであったという結果が得られている．しかし，キャプションを代替テキストとして再利用することは，ガイドライン [3] で挙げられている項目の 1 つ目，*“Do not duplicate float caption text as it detracts from the normal flow of your article and can, potentially, confuse the reader.”* に反するものであり，知識や経験のない人が適切な代替テキストを記述することの難しさを示している．

このような背景を踏まえ，代替テキストの記述を行う上での注意点や，代替テキストの質を向上させるために必要なことをインタビューなどを通して明らかにしようとする研究も盛んに行われている [1], [13], [14]．具体的には，グラフなどのデータを可視化した画像を代替テキストで表現する際の記述内容の詳細度をモデル化する研究 [14] や，代替テキストを記述する上での文脈の重要性を調査する研究 [13] などがある．また，Muehlbradt らの研究 [15] では，視覚障害のある方が本当に求めている情報と代替テキストの記述内容に乖離があるのではないかという疑問に対し，視覚障害のある方とその介助者が一緒に代替テキストを記述するワークショップを行った．そこでは，代替テキストを受け取る人と記述する人が特定の概念を共有していることの大切さや，スクリーンリーダーユーザーが代替テキスト作成に協力することの重要性などが報告された．

しかし，これらの研究は一般的な代替テキスト記述に対する多くの知見を提供する一方で，その多くは学術論文内の画像における代替テキスト記述に着目したものではない．論文内の画像は，代替テキスト研究において研究対象とされることの多い SNS や Web サイト上の画像と比較して，

- 画像の種類が，写真からグラフまで多岐にわたる．
- 文脈依存性が高く，内容が高度である．
- グラフ，スクリーンショット，概念図など，複雑だがその内容を確実に正しく伝えなければならないものが多い．

など, 独自の特性と記述の難しさを持つとされている [1], [5].

よって，学術論文内の画像に対して代替テキスト記述を行う上で何が大切か，著者は代替テキストを書く際にどのような間違いをしやすいのか，それに対してどういった支援が必要なのかを明らかにするためには，論文内画像に対する代替テキストの記述状況や記述内容に関する調査が必

要とされる．

2.3 代替テキスト記述支援への生成 AI の活用

近年の生成 AI の台頭により，代替テキストの生成及び記述支援においても，生成 AI を活用する試みが盛んに行われている [6], [7], [16]．その一例が，FigurA11y [6] である．このシステムでは，生成 AI によって代替テキストの初稿を作成し，その後利用者とのインタラクションを通して，代替テキストを修正していくことができる．このシステムは論文内の画像に対する代替テキストの記述支援を目的としている．また，SaGol [7] は，MiniGPT-4 を用いて画像の代替テキストを生成するシステムであり，スマートフォン内の画像に対して自動的に代替テキストを生成し，ユーザに提供することができる．

これらの研究のような，生成 AI の代替テキスト記述支援への活用は，今後ますます進展していくことが予想されるが，論文内画像の代替テキスト記述を支援するシステムを作る上で，生成 AI に入力するプロンプトの設計や，システムがどのようなインタラクションや機能を持つべきかといったことに関する知見はまだ十分とは言えない [8]．したがって，代替テキスト記述の現状を踏まえ，こういったシステムの設計に活用可能な知見を提供するような研究が必要とされている．

2.4 代替テキストのデータの収集・分析

これまでにも，代替テキストを収集した研究はいくつか行われており，例えば，Sharma ら [17] は，ウェブから収集したの大量の画像と，それに対する，高度に抽象化された代替テキストを提供することを目的とした研究を行った．この研究は，そのデータセットとしての利用価値から画像処理分野をはじめとした多くの研究で利用されている．

また，Crane ら [18] は，学術誌における図の代替テキストの設定状況について調査を行っており，主に医療分野の論文を対象に，代替テキストの設定状況を調査している．

そして，より HCI 分野の学術論文に焦点を当てた研究としては，Chintalapati ら [5] の研究と，Williams ら [19] の研究が挙げられる．

Chintalapati ら [5] の研究では，HCI 分野の国際会議である CHI と ASSETS において，2010 年から 2020 年までに発表された論文をランダムに選択し，その中から，データ可視化に関連する画像に設定されている代替テキストを収集した．この研究では合計 3,386 の，データ可視化に関連する画像と代替テキストが収集され，その記述の詳細度を Lundgard ら [14] の手法に基づいてラベリングした．

また，Williams ら [19] は，CHI 2019-2021 において発表された論文の中から 300 件の代替テキストを収集し，その内容を分析した．加えて，代替テキスト記述を行ったことのある人にユーザインタビューを行うことで，代替テキスト記述における困難を明らかにした．

これに対し本研究では，CHI 2021-2024 において発表された全ての論文（約 3,000 件）の中の全ての画像に対して，その代替テキストの内容について統計的な分析を行うことで，コミュニティが抱える課題を明らかにしつつ，今後，アクセシビリティに関する研究に利用可能な大規模なデータセットを提供することを目指す．本研究の貢献は大きく，

- 約 17,000 件という，先行研究と比較して大規模なデータの収集
- 大量なデータに対する ACM の提供するガイドライン [3] に則した分析
- CHI において，代替テキストの設定が義務化されて以降の [12] 近年のデータに対する調査

を行い，

- 現在の HCI 分野の学術論文における代替テキストの設定状況と記述傾向の調査結果
- 代替テキスト記述支援システムに活用可能な知見
- 今後の研究に利用可能な大規模なデータセット

を提供することにある．

3. データ収集

3.1 データの収集の概要

本研究では，ヒューマンコンピュータインタラクション分野の国際会議である CHI において，2021 年から 2024 年までに発表された論文を対象に，論文内の画像に設定されている代替テキストを収集した．その結果，3,319 本の論文から合計 24,306 枚の画像が収集された．

3.2 データ収集の詳細

Chintalapati ら [5] の研究では，ダウンロードした PDF ファイルを Adobe Acrobat で HTML に変換した上で代替テキストを取得していたが，本研究では，ACM Digital Library[*1]から HTML 形式で論文をダウンロードし，そこから，Figure タグを抽出する形で画像と代替テキストを抽出した．

具体的に収集した情報は以下の通りである．

- 論文の識別子（DOI）
- 会議名（今回は全て CHI になる）
- 発行年
- タイトル
- キーワードのリスト
- 代替テキストが設定された図の総数
- 図の総数
- 論文内に含まれる図の情報
  - 画像ファイルの URL
  - キャプション

[*1] https://dl.acm.org/

– 代替テキスト（存在しない可能性もある）

なお，論文のダウンロードを行う際には，ACM Digital Library に負荷をかけないように，十分な間隔を空けてダウンロードを行った．

4. 結果 (Results)

4.1 データの収集結果 (Data Collection Results)

表 1 に，収集できたデータの概要を示す．今回のデータ収集では，合計 3,319 本の論文から 24,306 枚の画像が収集できた．そのうち 17,721 枚の画像に代替テキストが設定されていた．

表 1 CHI 2021-2024 における代替テキストの設定状況

年度	図を持つ論文 / 全論文	代替テキストを持つ図 / 全ての図
2021	681 / 746 (91.3%)	3565 / 4778 (74.6%)
2022	599 / 637 (94.0%)	3209 / 4630 (69.3%)
2023	834 / 879 (94.9%)	5288 / 6713 (78.8%)
2024	1008 / 1057 (95.4%)	5659 / 8185 (69.1%)
合計	3122 / 3319 (94.1%)	17721 / 24306 (72.9%)

4.2 データ分析

ここで収集されたデータに対して，ACM の提供するガイドライン [3] の項目に基づいて分析を行った．ただし，代替テキストが設定されている画像（17,721 件）のみを分析対象とした．

4.2.1 代替テキストの長さ

代替テキストの文字数の分布を図 1 に示す．ただし，代替テキストが長すぎるものがあるので，図においては上位 1%を除外して表示している．

また，図中の赤い線は，ACM の提供するガイドライン [3] の項目 3（*“Keep it (relatively) short. The most popular screen readers cut off alt text at around 125 characters, so it’s advisable to keep it to that character count or less.”*）において推奨されている 125 文字を示している．約 80%のデータがこの文字数を超えてしまっていることがわかる．

4.2.2 代替テキストとキャプションの関係

代替テキストとキャプションの関係については，ガイドライン [3] の項目 1（*“Do not duplicate float caption text as it detracts from the normal flow of your article and can, potentially, confuse the reader.”*）を踏まえ，ここでは，代替テキストとキャプションの同一性，内容の類似度についての分析を行った．

まず，代替テキストがキャプションと同一，すなわち (キャプション) = Figure: + (図の番号) + (代替テキスト) となるものがいくつ存在するかを調査した．その結果，17,721 件のうち，959 件（5.4%）が代替テキストとキャプションが同一であることが分かった．

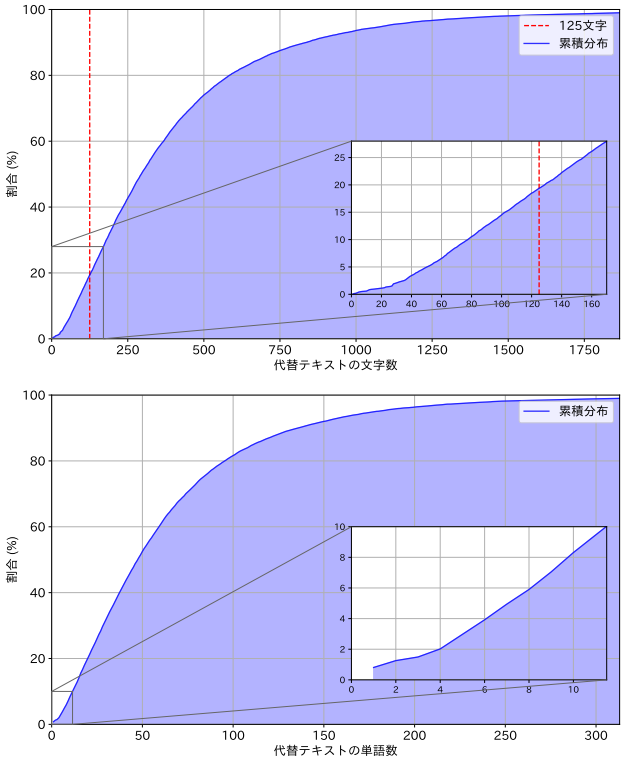

図 1 代替テキストの文字数と単語数の累積分布．一部の代替テキストが長すぎるので，図では上位 1%は除外した．赤い線はガイドラインで推奨されている 125 文字を示している．

Fig. 1 Cumulative distribution of alternative text lengths and word counts. The red line indicates the recommended 125 characters in the guideline.

次に，代替テキストとキャプションについて，その長さの関係を可視化した．その結果を図 2 に示す．分布からキャプションよりも代替テキストの方が少しだけ単語数が多い傾向が見て取れる．また，図中，点群により傾き 1 の直線が出現しているが，この直線は $y = x + 2$ で，代替テキストとキャプションが同一であるものが分布する直線である．

次に，代替テキストの記述内容自体の類似性を調査するために，代替テキストとキャプションのレーベンシュタイン距離 [20] による類似度を計算した．その結果を図 3 に示す．この類似度は，0 から 1 の間に分布し，値が大きいほどキャプションと代替テキストが類似していることを示すが，この図においては，0.4-0.5 あたりの最大のピークと 1.0 近辺のピークという二つのピークが見られることが分かる．

4.2.3 *image of* や *picture of* を含む代替テキスト

最後に，ガイドラインの項目 6（*“Don’t include “image of,” “picture of,” etc. in your alt text. It’s already assumed your alt text is referring to a float, so there’s no need to specify it.”*）に基づいて，代替テキストの中に冗長な表現である「image of」や「picture of」が含まれているものがどれだけ存在するかを調査した．

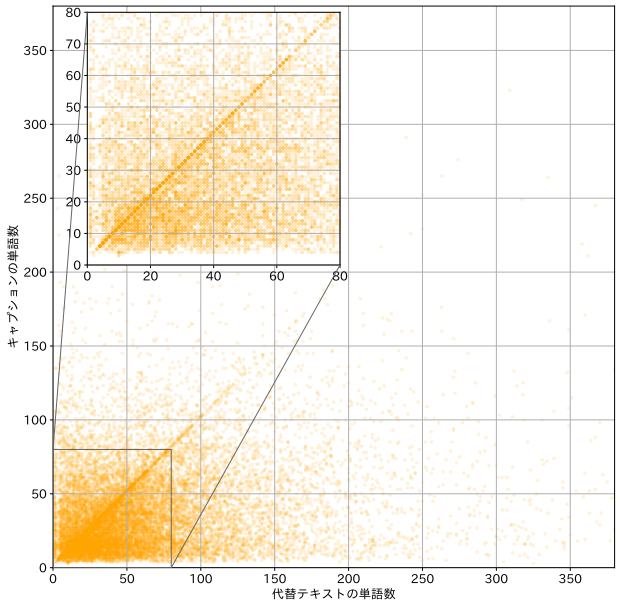

図 2 代替テキストとキャプションの単語数の比較．代替テキストの長さについては上位 1%を除外して表示している．図中，点群により傾き 1 の直線（$y = x + 2$）が出現しているが，これは代替テキストとキャプションの記述内容が同じものが分布するところである．

Fig. 2 Comparison of the number of words in alternative texts and captions.

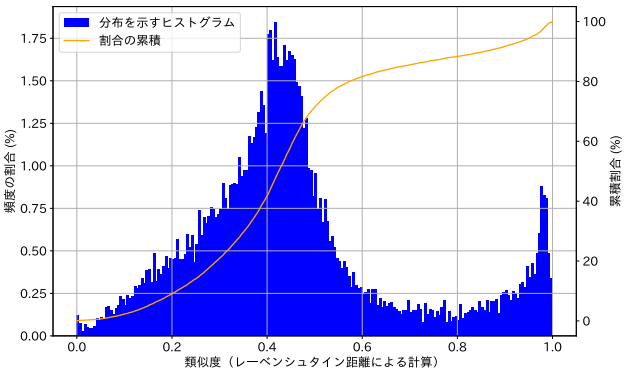

図 3 代替テキストとキャプションの間の類似度について，レーベンシュタイン距離 [20] を用いて計算し，その分布を示したもの．ヒストグラムは各類似度の出現割合を示しており，曲線はその累積分布を示している．

Fig. 3 Distribution of Levenshtein distances between alternative texts and captions. The histogram shows the occurrence rate of each similarity, and the curve shows the cumulative distribution.

その結果，「picture of」を含むものは 176 件（1.0%），「image of」を含むものは 379 件（2.1%）存在した．そして，その両方を含むものは 12 件（0.1%），いずれかを含むものは 543 件（3.1%）存在することが分かった．

5. 考察

今回得られた結果を踏まえ，現状の代替テキストの記述においてどのような課題や傾向があるのかについて考察する．

5.1 代替テキストの設定状況

まず初めに，代替テキストの設定状況については，表 1 からも分かる通り，全体の 72.9%の画像に代替テキストが設定されていた．これは，Chintalapati ら [5] の研究においては 19,500 件の論文のうち，有効な代替テキストを少なくとも一つでも持つものが 897 件（4.6%）であったことと比較すると，かなり高い割合であると考えられる．（ただし，Chintalapati らの代替テキストの抽出方法と，我々の用いた代替テキストの抽出方法は異なるため，直接の比較は難しい．）このような結果が得られたのは，CHI への論文投稿時に，代替テキストの設定が（少なくとも 2022 年以降は）実質義務付けられるようになった [12] ことによる影響が大きいと考えられる．しかし，それでも全体の 27.1%の画像には代替テキストが設定されていないということは，まだまだアクセシビリティの観点から見て改善の余地があると言えるだろう．

また，今回調査対象となった 4 年分のデータについて，年度ごとの代替テキストの設定状況についてはあまり大きな差は見られなかった．

5.2 代替テキストの長さ

次に，代替テキストの長さについて考えていく．

5.2.1 代替テキストが極端に長いもの

今回収集されたデータにおいて調査を行った結果，いくつか代替テキストの中には極端に長いものが存在することが分かった．最も長かったものは，3512 単語，242269 文字のもので，この画像が含まれる論文では平均 706 単語，4477 文字の代替テキストが設定されていた．そして，このような極端に長い代替テキストを持つものは，その多くが複数のグラフや画像が一つの画像としてまとめられているような複雑な図であった．これらの代替テキストは画像の内容を正しく伝えようとする意図があるものの，その長さから，代替テキストとしては適切でないと言えるだろう．

5.2.2 代替テキストが極端に短いもの

その一方で，極端に短い代替テキストは，図 1 の原点近辺の拡大図からもわかるようにかなり多く存在しており，10 単語以下の代替テキストだけで全体の約 8%にまで及んでいる．特に，明らかに不適切であるといえる，1 単語のみで構成される代替テキストが 17,721 件中，143 件（0.8%）存在しているのは問題であると言えるだろう．

この結果を踏まえ，これらの極端に短い代替テキストは

具体的に何が記述されているのかの調査を行った．そこで見つかった不適切な代替テキストのパターンを表 2 に示す．

表 2 極端に短く，不適切な代替テキストの例

内容	例
代替テキストとして意味を持たないもの	Description, A, Teaser, xx, long description
理解できない文字列	n.kjbkj, 16jietu
インデックスがついた無意味なもの	Figure 1, Fig 2
後で記入する印がそのままになっているもの	TODO, TBD, TBW, insert description here
英語でない文字列	「画像」の中国語
PC におけるパス	../../Desktop/Screen%20Shot%202022-03-08%20at%10.48.06%20PM.png
説明を他の場所に任せるもの	”As described in the caption.”, ”Fully described in the text.”
画像の種類を一語で表しただけのもの	plots, table, application, screenshot of XXX

このような明らかに不適切な代替テキストが少なくない数存在することは，コミュニティにとって大きな問題であると言える．現在の CHI においては，提出時に代替テキストが要求されるとされているが [12]，何かしらの文字列を入力すれば良いため，アクセシビリティについての関心がない著者がこのような不適切な代替テキストを設定してしまっていると予想される．

この課題を解決するには，例えば提出する代替テキストの長さに最低値を設けるといったアプローチが考えられるが，ただ，代替テキスト記述において短いことが必ずしも悪いというわけではない上に，記述意欲のない著者には簡単に突破されてしまうことが予想されるため，単純な文字数制限の効果は限定的である．よって，この課題を解決するには，論文の著者の意識・意欲の向上が不可欠である．そのためには適切なガイドラインや支援ツールの提供などを通した適切な代替テキストを記述することの心理的負担の軽減や，記述内容を加味したバリデーションなどが必要であると言えるだろう．

5.2.3 全体の傾向とガイドラインとの比較

前節までの極端な例について着目したが，ここでは全体の代替テキストの長さについて考えていく．ガイドライン [3] の項目 3 では，代替テキストは 125 文字程度もしくはそれ以下であることが推奨されているが，代替テキストの長さの分布については，図 1 を見ると，全体の約 80%が 125 文字を超えてしまっており，ガイドラインで推奨されている範囲内に収まっているものは少ないことが分かる．

そして，Williams ら [19] の研究では，同じく CHI における代替テキストの記述内容について，その説明の詳細度について評価を行ったところ，代替テキストが 40 単語を超えたあたりからある程度説明的になりはじめ，70 単語以上になると内容が十分説明的である傾向が強まるという結果が報告されている．加えて，英語の文字数について，一般的な英語の一単語あたりの文字数は 3 文字が最頻値で，平均はそれより少し大きい [21] ことが報告されている．

これらを踏まえると，良質な代替テキストを記述しようとする中で，その文字数を 125 文字以内に収めるという制約は，必要以上に厳しい制約となる可能性が示唆される．加えて，ガイドラインで推奨されている 125 文字程度という文字数は，スクリーンリーダーの一般的な設定という，実装次第で変更可能な制約に基づくものであるため，代替テキスト自体の質という観点からするとあまり本質的ではない．もちろん，あまりにも長すぎる代替テキストは適切ではないので，文字数に対する制約は必要であるだろうが，このガイドラインの適切性については，今後の検証が必要であると考える．

5.3 代替テキストとキャプションの関係

次に，ガイドライン [3] の項目 1 では，代替テキストとキャプションが重複していることは避けるべきでると述べられていることを踏まえ，代替テキストとキャプションの関係について考えていく．

5.3.1 代替テキストとキャプションの単語数の比較

今回得られたデータにおける代替テキストとキャプションの単語数の分布を図 2 に示した．この図を見ると，代替テキストの単語数は，キャプションの単語数よりもやや多い傾向が見られる．なお，単語数の最大値についても，キャプションが 422 単語であるのに対し，代替テキストが 3512 単語と大きな差があった．キャプションに対し，代替テキストが長くなりがちな理由としては様々な理由が考えられるが，その理由の一つとして，キャプションは論文ファイルの中で視覚的に表示されるものであるため，スペースの制約があるのに対し，代替テキストは画像の属性として設定されるものであるため，その長さに制約がないことが予想される．この，現状の代替テキスト記述を行う体制では長さの制約が発生せず，その長さが著者に依存してしまっていることは，前節の内容とも関連して，支援ツールの開発などを行う上で重要な観点になるだろう．

また，図 2 において着目すべきは，原点近辺から傾き 1 の直線に沿って分布している点が多く見られることである．今回のデータにおいて，代替テキストとキャプションの内容が同一であるものは，5.4%存在することが分かったが，それらのデータは，この図において $y = x + 2$ の直線上に位置する．先行研究 [1] においては 60%の代替テキストがキャプションと同一であると報告されていたため，そ

の結果と比較すると，その割合はかなり低いといえるが，依然として散布図において確認できるほど，ガイドラインに抵触するような代替テキストとキャプションの関係が存在していることは非常に重要な知見である．

5.3.2 代替テキストとキャプションの内容の類似度

単純な代替テキストとキャプションの一致判定では，代替テキストとキャプションの内容の類似度を正確に評価することは難しいので，代替テキストとキャプションの内容の類似度をレーベンシュタイン距離を用いて評価した．その結果が図 3 であるが，この図において興味深い点は，類似度の分布のピークが 2 つ存在することである．代替テキストとキャプションは，内容的に独立であるわけではないので，何かしら類似度が高い傾向があることは予想されるが，類似度 1.0 近辺にこれほどのピークが存在することは特筆すべきである．このピークが存在することから，代替テキストを書く際に，代替テキストとしてキャプションをコピーアンドペーストする形だけではなく，雛形としてキャプションを利用して，そこから少し変更を加えるといった形で代替テキストを記述すること人が一定数いることが予想される．適切な改変を加えるのであれば，文脈に基づいた雛形が提供される分，一から代替テキストを記述するよりも容易であるため，一つの良い戦略となり得るが，その改変が適切であるかどうかを判断するための支援ツールが必要であると言えるだろう．もしくは，FigurA11y [6] のように，画像の内容に基づいた代替テキストの雛形を提供することで，利用者の負担を軽減しながら，適切な代替テキストの質も向上させるといったアプローチも考えられるだろう．

5.4 *image of* や *picture of* を含む代替テキスト

次に，ガイドライン [3] の項目 6 では，代替テキストに冗長な表現である「image of」や「picture of」が含まれていることは避けるべきであるが，今回のデータにおいては，そのような表現が含まれている代替テキストが 543 件（3.1%）存在していた．この項目のように，ルールベースで判定可能な簡単な問題については，代替テキスト記述支援システムに，バリデーションを行う機能を実装するなどして対処することができるだろう．

5.5 既存の代替テキスト自動生成ツールの利用状況

最後に，現状の代替テキストにおいて，代替テキスト自動生成ツールがどのように利用されているのかについて考察する．

今回収集されたデータにおいて，「A graph of blue and white bars Description automatically generated」,「Graphical user interface, text, application Description automatically generated」など，代替テキストが「Description automatically generated」で終了するものが合計 80 件（0.5%）存在した．これらの代替テキストは，何かしらの代替テキスト自動生成ツールを利用して生成されたものであると考えられる．このサフィックスが付与されているものは，抽象的かつ十分な情報を提供していないため，論文内の画像に対する代替テキストとしては不十分なものが多いが，代替テキスト記述支援システムの中で，自動生成機能のようなものを提供する価値を示唆している．

加えて，画像を解釈できるマルチモーダルな生成 AI が広く普及したことを踏まえると，今後，生成 AI を代替テキスト作成に利用する著者は今よりも増えていくであろうし，今回収集されたデータにおいても，生成 AI によって記述された代替テキストが含まれている可能性は高いと考えられる．よって，論文査読における生成 AI の利用状況を査読データから推定した Liang [22] の研究のように，代替テキストにおける生成 AI の利用状況について調査を行うことは，今後の研究の展望として考えられる．

6. 終わりに

本研究では，近年のアクセシビリティ向上意識の高まりを受け，学術論文における代替テキスト設定の状況とその記述内容の傾向について調査を行うために，近年の CHI の論文データから 17,721 件の代替テキストを収集し，そのデータに対する分析を行った．

その結果，代替テキストの設定状況や長さ，代替テキストとキャプションの関係についての現状を明らかにし，代替テキスト記述支援システムやガイドラインの改善に利用可能な知見を得ることができた．しかし，ガイドラインの中でも，ルールベースでは判断が難しい内容についての分析を行えていない点や，他にも多数存在するガイドラインについての研究を行えていない点は本研究の課題である．

今後の研究の展望としては，以上の課題を解決するとともに，アクセシビリティ関連の研究において利用可能な代替テキストのデータセット提供や，得られた知見に基づく代替テキスト記述支援システムの開発を行っていきたいと考えている．

参考文献

[1] Bigham, J. P., Brady, E. L., Gleason, C., Guo, A. and Shamma, D. A.: An uninteresting tour through why our research papers aren't accessible, *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems*, pp. 621–631 (2016).

[2] Mack, K., McDonnell, E., Jain, D., Lu Wang, L., E. Froehlich, J. and Findlater, L.: What do we mean by "accessibility research"? A literature survey of accessibility papers in CHI and ASSETS from 1994 to 2019, *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–18 (2021).

[3] ACM: How to Write Alt Text and Why, https://authors.acm.org/journals/how-to-write-alt-text-and-why (2021).

[4] SIGACCESS: Describing Figures, https://www.sigaccess.org/welcome-to-sigaccess/resources/describing-figures/ (2020).
[5] Chintalapati, S. S., Bragg, J. and Wang, L. L.: A dataset of alt texts from HCI publications: Analyses and uses towards producing more descriptive alt texts of data visualizations in scientific papers, *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 1–12 (2022).
[6] Singh, N., Wang, L. L. and Bragg, J.: FigurA11y: AI Assistance for Writing Scientific Alt Text, *Proceedings of the 29th International Conference on Intelligent User Interfaces*, pp. 886–906 (2024).
[7] Moon, Y., Lee, H., Oh, S. and Jung, H.: SaGol: Using MiniGPT-4 to Generate Alt Text for Improving Image Accessibility.
[8] Mack, K., Cutrell, E., Lee, B. and Morris, M. R.: Designing tools for high-quality alt text authoring, *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 1–14 (2021).
[9] Bourne, R., Steinmetz, J. D., Flaxman, S., Briant, P. S., Taylor, H. R., Resnikoff, S., Casson, R. J., Abdoli, A., Abu-Gharbieh, E., Afshin, A. et al.: Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study, *The Lancet global health*, Vol. 9, No. 2, pp. e130–e143 (2021).
[10] W3C Web Accessibility Initiative: Web Accessibility Tutorials: Guidance on how to create websites that meet WCAG, https://www.w3.org/WAI/tutorials/images/ (2019).
[11] SIGCHI: Accessibility Guides for Authors, https://sigchi.org/resources/guides-for-authors/accessibility/ (2022).
[12] CHI 2022: Publication-Ready Author Instructions, https://chi2022.acm.org/for-authors/publication-ready-author-instructions (2022).
[13] Kreiss, E., Bennett, C., Hooshmand, S., Zelikman, E., Morris, M. R. and Potts, C.: Context Matters for Image Descriptions for Accessibility: Challenges for Referenceless Evaluation Metrics, *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pp. 4685–4697 (2022).
[14] Lundgard, A. and Satyanarayan, A.: Accessible Visualization via Natural Language Descriptions: A Four-Level Model of Semantic Content, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 28, No. 1, pp. 1073–1083 (online), DOI: 10.1109/TVCG.2021.3114770 (2022).
[15] Muehlbradt, A. and Kane, S. K.: What's in an ALT Tag? Exploring Caption Content Priorities through Collaborative Captioning, *ACM Transactions on Accessible Computing (TACCESS)*, Vol. 15, No. 1, pp. 1–32 (2022).
[16] Leotta, M., Mori, F. and Ribaudo, M.: Evaluating the effectiveness of automatic image captioning for web accessibility, *Universal access in the information society*, Vol. 22, No. 4, pp. 1293–1313 (2023).
[17] Sharma, P., Ding, N., Goodman, S. and Soricut, R.: Conceptual captions: A cleaned, hypernymed, image alt-text dataset for automatic image captioning, *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 2556–2565 (2018).
[18] Crane, M. A., Nguyen, M., Lam, A., Berger, Z. D., Paulus, Y. M., Romley, J. A. and Faden, R. R.: Figure accessibility in journals: analysis of alt-text in 2021–23, *The Lancet*, Vol. 402, No. 10419, pp. 2287–2289 (2023).
[19] Williams, C., de Greef, L., Harris III, E., Findlater, L., Pavel, A. and Bennett, C.: Toward supporting quality alt text in computing publications, *Proceedings of the 19th International Web for All Conference*, pp. 1–12 (2022).
[20] Gusfield, D.: *Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology*, EBL-Schweitzer, Cambridge University Press (1997).
[21] Sigurd, B., Eeg-Olofsson, M. and Van Weijer, J.: Word length, sentence length and frequency–Zipf revisited, *Studia linguistica*, Vol. 58, No. 1, pp. 37–52 (2004).
[22] Liang, W., Izzo, Z., Zhang, Y., Lepp, H., Cao, H., Zhao, X., Chen, L., Ye, H., Liu, S., Huang, Z. et al.: Monitoring ai-modified content at scale: A case study on the impact of chatgpt on ai conference peer reviews, *arXiv preprint arXiv:2403.07183* (2024).