

Paper2Poster: LLM を用いた学術ポスターの生成

増田 大河^{1,2} 田中 翔平² 平澤 寅庄² 牛久 祥孝²
¹ 中部大学 ² オムロンサイニックス株式会社

masudataiga@mprg.cs.chubu.ac.jp
{shohei.tanaka, tosho.hirasawa, yoshitaka.ushiku}@sinicx.com

概要

学術ポスターは文章と図表を組み合わせる研究内容をまとめたポスターであり、論文の概要を短時間で理解するのに効果的なフォーマットである。従来の学術ポスターの自動生成に関する研究は主にレイアウトの生成を目的としており、コンテンツも含めた生成は行っていない。そこで本研究では、学術ポスターに含まれる文章や図表、レイアウトを LLM を用いて段階的に予測することで、学術ポスター全体を生成する Paper2Poster システムを構築した。定量的および定性的な評価結果より、構築したシステムは整列されたオーバーラップのない、視覚性に優れたポスターを生成可能であることがわかった。

1 はじめに

学術論文は研究成果がまとまったものとしては基本的なフォーマットだが、先行研究の論文を精読して内容を把握することは時間のかかる作業である。これに対して学術ポスターは文章や図表を用いて研究内容をまとめたポスターであり、短時間で研究の概要を把握するのに効果的なフォーマットである。しかし学術ポスターは学会の開催中に現地のみで公開され、Web 上で公開されないことも多い。また arXiv やジャーナルなどで公開されている学会に紐づかない論文の場合、そもそもポスターが存在しないことがほとんどである。だが先行研究をまとめたポスターを手で作成するには論文を精読する必要があり、学術ポスターの目的である研究概要の把握の高速化が達成できない。ここで論文から学術ポスターを自動で生成するシステムを構築することができれば、先行研究を把握する作業を高速化し、研究活動を促進することができると考えられる。

学術ポスターのためのレイアウト認識や生成を目的とした SciPostLayout [1] は、LLM を用いた手法が

論文から学術ポスターのレイアウトを生成できる可能性を示している。しかし SciPostLayout における実験はコンテンツを考慮しないレイアウト生成にとどまっており、ポスターのコンテンツも含めたレイアウト生成およびポスター生成は行っていない。これに対して人間が学術ポスターを作成する場合、セクションや図表の数だけではなく、各要素に含めるコンテンツも考慮してレイアウトを決定する。

そこで本研究では、LLM による多段階推論を用いてレイアウトとコンテンツの両方を考慮した学術ポスターを生成する Paper2Poster システムを提案する。具体的には、学術ポスターに含まれる文章や図表を LLM を用いて段階的に生成し、その条件に基づいてレイアウトを生成することで完成された学術ポスターを生成するシステムを構築した。構築したシステムを評価するため、正解のポスターとの類似度などに基づく定量的評価、および人間の研究者による定性的評価を実施した。実験結果より、提案するシステムは整列されたオーバーラップのないレイアウトを生成可能であり、視覚性や構造的な主観的な評価項目においても優れたポスターを生成可能であることがわかった。

2 関連研究

論文から図表を用いた資料を作成する関連研究としては、Bandyopadhyay [2] らによるプレゼンテーションスライドの自動生成システムがある。Bandyopadhyay らは以下の3つの理由からオープンメインな生成タスクと比較してスライド生成タスクをより難しいタスクだと定義している。(1) 長い文書全体を LLM に与えることは、コンテキスト長制限のために困難である [3]。(2) LLM の性能はプロンプト内のコンテキストの長さに比例して低下する [4]。(3) 生成されたコンテンツの正確な配置が苦手であり、ハルシネーションを起こしやすい [5, 6]。

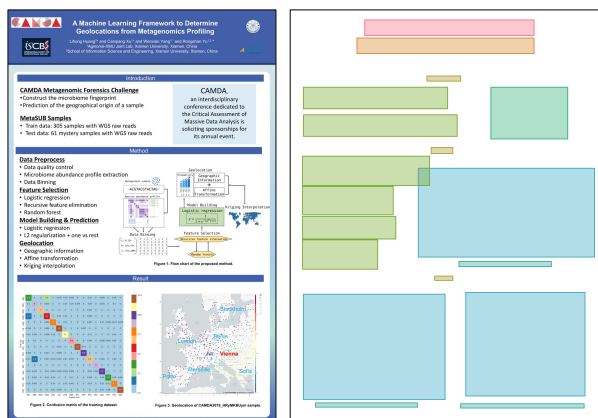


図 1 SciPostLayout データセットの例：収録されている学術ポスター（左）およびアノテーションされたカテゴリと BBox 情報（右）

これらの理由から、Bandyopadhyay らは論文からスライドを直接生成するのではなく、スライド生成を単純な複数のサブタスクに分割し、これらを段階的に処理することでプレゼンテーションスライドを生成するシステムを実現している。本研究ではこのシステムを参考に、多段階で論文から学術ポスターを生成するシステムを構築した。

田中ら [1] は学術ポスターのレイアウト認識や生成を目的として、手動でレイアウトがアノテーションされた 7,855 枚の学術ポスターとポスターに紐づけられた 100 本の論文が収録されている SciPostLayout データセットを構築した。SciPostLayout データセットの例を図 1 に示す。このデータセットにおけるアノテーションカテゴリは Title, Author Info, Section, Text, List, Table, Figure, Caption, Unknown の 9 種類である。構築したデータセットを用いて、田中らは学術ポスターのレイアウト生成実験を行なった。ここで論文からポスターレイアウトを生成するタスクとして、Gen-T と Gen-P の 2 つのタスクが定義されている。Gen-T では GPT-4 を用いて論文から抽出された要素タイプ制約に基づきレイアウト生成モデルがレイアウトを生成する。Gen-P では GPT-4 を用いて論文を 1,000 単語以内に要約し、この要約に基づいてレイアウト生成モデルがレイアウトを生成する。これらのタスクにおけるレイアウト生成実験では、LLM ベースのモデルが整列されたオーバーラップの少ないレイアウトを生成可能であり、LLM が論文から学術ポスターを生成するシステムとしてのポテンシャルを秘めていることが確認された。本研究では SciPostLayout のレイアウト生成フレームワークを発展させ、ポス

ターのコンテンツも含めた生成を行うシステムを構築することを目指す。

3 提案手法

SciPostLayout の Gen-T と Gen-P の 2 つのタスクは、要素タイプ制約と要約文といった簡潔な条件でのレイアウト生成であり、各セクションに含まれる文章の長さや図表の数といった詳細な条件は考慮していない。そのため、生成されたレイアウトを用いてコンテンツも含んだ学術ポスターを直接生成することは困難である。そこで、論文から学術ポスターを生成するプロセスを以下の 3 つのサブタスクに分割して学術ポスターを生成する Paper2Poster システムを提案する。提案システムの処理手順は以下の 3 つに分割される。(1) 学術ポスターに含むコンテンツ（文章や図）の生成。(2) 学術ポスターのレイアウトの生成。(3) 生成したコンテンツとレイアウトから学術ポスターを生成。これらの処理を段階的に実施することで、学術ポスターに含める具体的なコンテンツに基づいたレイアウトを生成することができる。また生成したレイアウトと文章や図表のコンテンツを組み合わせることで完成された学術ポスターを生成することが可能となる。本手法は In-context Learning を活用して既存の LLM のパラメータの調整を行うことなく学術ポスターの自動生成を実現している。Paper2Poster システムのフレームワークの概要を図 2 に示す。また LLM を用いた処理に使用したプロンプトを付録 A.1 に示す。

学術ポスター生成を目的とした要約文の生成

始めに、LLM を用いて学術ポスター生成を目的とした論文の要約を行う。これにより、論文に含まれる大量のテキストから学術ポスターに含めるべき重要な部分を抜き出す。このとき、学術ポスターに含めるべき図や表も抜き出すように指示する。ここで LLM は基本的にテキストのみを処理するためのモデルであり、図や表を画像として処理することができない。そのため、図や表のキャプションや本文の引用部分に着目して学術ポスターに含めるべきか否かを判断するようにプロンプトを設計した。

構造化データの生成

次に要約された論文テキストから LLM を用いて学術ポスターを生成するための構造化データを生成する。ここでいう構造化データとは、どのセクションにどの文章や図表が含まれるかを対応付けた構造化データである。ここでは zero-shot でプロンプトにタ

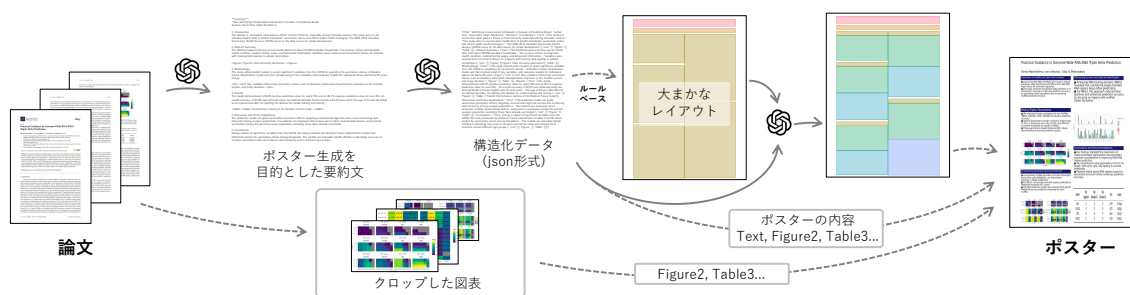


図2 Paper2Poster のフレームワークの概略図

スクと出力フォーマットを明示することで期待する出力を獲得している。

レイアウトの生成

ここでは構造化データを条件として2段階でレイアウトを生成する。まず1段階目の生成ではルールベースにより大まかなレイアウトを生成する。このとき学術ポスターの向きや列数の条件に基づいたレイアウト生成を行う。次に2段階目の生成では、1段階目で生成したレイアウトと構造化データを条件とする詳細なレイアウトをLLMを用いて生成する。

学術ポスターの生成

最後にコンテンツを含む構造化データとレイアウトから学術ポスターを生成する。提案システムでは図表の生成は行わず、論文からクロップした図表の画像データをそのままポスターに配置する。具体的には生成したレイアウトの各領域と構造化データの各コンテンツをマッチングさせ、各領域内に文章や図表を配置することで、最終的な画像データとして学術ポスターを生成する。

4 実験

提案システムの性能を評価するために、生成された学術ポスターのコンテンツとレイアウトを自動評価手法を用いて定量的に評価した。また生成された学術ポスターが論文の内容をわかりやすくまとめているかを人手評価を用いて定性的に評価した。生成された学術ポスターのコンテンツとレイアウトの定量的評価では、SciPostLayout データセットのテストデータに含まれる50本の論文とポスターのペアを使用した。生成された学術ポスターの定性的評価ではWeb上で収集したCVとNLP分野の論文を使用した。システムを構成するLLMとして、GPT-4o¹⁾の2つスナップショット(2024-05-13, 2024-08-16)を使用した。

1) <https://openai.com/index/hello-gpt-4o/>

表1 コンテンツ生成の定量的評価結果(コサイン類似度)。太字は最良の結果を表す。

Model	Sentence-Sim	Section-Sim
gpt-4o-2024-05-13	0.455	0.611
gpt-4o-2024-08-16	0.451	0.640

4.1 コンテンツ生成の定量的評価

生成された学術ポスターに含まれる文章と実際の学術ポスターの文章との類似度について、文埋め込みモデル[7]を用いて変換した埋め込みベクトルのコサイン類似度で評価した。評価指標として以下の2つの指標を用いた。(1) Sentence-Sim: 文章を1文ずつに分割して実際のポスターの各文と類似度を計算し、最も高い類似度を評価値として使用する。(2) Section-Sim: 各セクションの文章ごとに実際のポスターの文章と類似度を計算し、最も高い類似度を評価値として使用する。評価結果を表1に示す。表より、Section-Simについてgpt-4o-2024-08-16が0.640のコサイン類似度を達成しており、実際の学術ポスターとある程度意味的に近い文章を生成できていると考えられる。またSentence-SimはSection-Simと比較して低い評価値となっているが、これは文単位で類似度を評価した場合の方が実際の文と類似していない文が含まれていた場合の悪影響が大きいためだと考えられる。

4.2 レイアウト生成の定量的評価

ここでは実際のレイアウトを必要としない指標としてalignment, overlapの2つの評価指標を用いた。Alignmentは生成されたレイアウト内の要素がどの程度整列されているかを評価する指標である[8, 9]。Overlapは任意の2要素についてオーバーラップしている領域の大きさを評価する指標である[8, 9]。どちらの指標も値が小さいほど良いレイアウトであることを意味する。また実際のレイ

表 2 レイアウト生成の定量的評価結果. Portlait n column は縦向き n 列のポスターを表し, input orientation は実際のポスターと同じ向きの 3 列のポスターを表す. ↑は値が大きいほど性能が高いことを意味し, ↓は値が小さいほど性能が高いことを意味する. 太字は最良の結果を表し, 下線は次点の結果を表す.

Task	Model	Alignment ↓	Overlap ↓	LTSim ↑
Gen-T	gpt-4o-2024-08-06	0.000	0.116	0.6859
Gen-P	gpt-4o-2024-08-06	0.000	0.021	0.6837
Paper2Poster (portrait 1 column)	gpt-4o-2024-05-13	0.000	0.000	0.6306
	gpt-4o-2024-08-06	0.000	0.000	0.6404
Paper2Poster (portrait 2 column)	gpt-4o-2024-05-13	0.000	0.000	0.6363
	gpt-4o-2024-08-06	0.000	0.000	0.6317
Paper2Poster (portrait 3 column)	gpt-4o-2024-05-13	0.000	0.000	0.6374
	gpt-4o-2024-08-06	0.000	0.000	0.6308
Paper2Poster (input orientation)	gpt-4o-2024-05-13	0.000	0.000	0.6338
	gpt-4o-2024-08-06	0.000	0.000	<u>0.6390</u>

表 3 ポスター生成の評価結果

	視覚性	構造的性	明瞭性	網羅性	正確性
平均値	3.80	3.78	3.70	3.43	3.75
標準偏差	0.65	0.80	0.79	0.90	0.93
最大/最小	5 / 2	5 / 2	5 / 2	5 / 1	5 / 2

アウトとの類似度を測る評価指標として, 最適輸送ベースの類似度評価指標である LTSim [10] を用いた. 評価結果を表 2 に示す. 表より, 提案する Paper2Poster システムは alignment, overlap がほぼ 0 のレイアウトを生成できていることが確認できる. また LTSim では Gen-T が最も高い性能を達成しているが, Paper2Poster も同等の性能を達成していることが確認できる.

4.3 ポスター生成の定性的評価

ここでは Paper2Poster (gpt-4o-2024-08-06) が生成した学術ポスターを定性的に評価した. SciPostLayout データセットに含まれる論文の分野は多岐にわたり, 生成されたポスターの適切な評価者を選定することが難しい. そこで本実験では CV 分野 5 本, NLP 分野 5 本の最新論文を新たに Web 上で収集し, これらの論文から学術ポスターを生成して定性的に評価した. 生成した 10 本の学術ポスターを視覚性, 構造的性, 明瞭性, 網羅性, 正確性の 5 つの観点についてそれぞれ 5 段階で評価した. 各評価指標の詳細は付録 A.2 に示す. 評価者として CV および NLP 分野に精通した 4 名の研究者を選定し, 評価を依頼した. 評価結果を表 3 に示す. 表より, 全体的に視覚性, 構造的性, 明瞭性の評価値が高いことから, ポスターのコンテンツが見やすいレイアウトで論理的に配置されていること, また情報が簡潔でありながらも要点が伝わるコンテンツが生成できていることが確認できる. Paper2Poster が生成した学術ポ

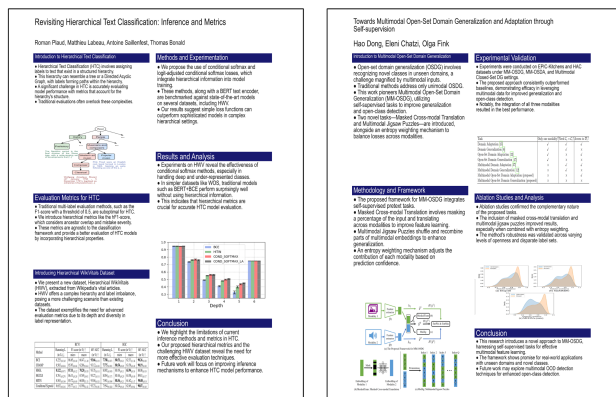


図 3 Paper2Poster が生成した学術ポスター

ターの例を図 3 に示す. 図より, 生成した学術ポスターのレイアウト要素が論理的に配置されており, 人手で作成する場合に近いポスターが生成できていることがわかる. しかし生成された学術ポスターの中には図の空白部分が多い場合やアスペクト比を無視した配置になっている場合も確認できる.

5 おわりに

本研究では学術ポスターを自動で生成する Paper2Poster システムを提案した. 提案システムはポスター生成を複数のサブタスクに分割し, 多段階でコンテンツやレイアウトを生成することで学術ポスターを生成する. 定量的な評価実験より, 提案システムは整列されたオーバーラップのないレイアウトを生成することが確認できた. また定性的な評価実験より, 情報が簡潔でありながらも要点が伝わるコンテンツが見やすいレイアウトで論理的に配置されている学術ポスターが生成できていることを確認した. 今後は大規模な Paper2Poster データセットを構築し, また学術ポスター生成システムの性能を向上させる方法について検討していく.

謝辞

本研究は、JST【ムーンショット型研究開発事業】【JPMJMS2236】の支援を受けたものです。

参考文献

- [1] Shohei Tanaka, Hao Wang, and Yoshitaka Ushiku. Scipost-layout: A dataset for layout analysis and layout generation of scientific posters, 2024.
- [2] Sambaran Bandyopadhyay, Himanshu Maheshwari, Anandhavelu Natarajan, and Apoorv Saxena. Enhancing presentation slide generation by LLMs with a multi-staged end-to-end approach. In Saad Mahamood, Nguyen Le Minh, and Daphne Ippolito, editors, **Proceedings of the 17th International Natural Language Generation Conference**, pages 222–229, Tokyo, Japan, September 2024. Association for Computational Linguistics.
- [3] Jesse Mu, Xiang Li, and Noah Goodman. Learning to compress prompts with gist tokens. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, **Advances in Neural Information Processing Systems**, volume 36, pages 19327–19352. Curran Associates, Inc., 2023.
- [4] Nelson F. Liu, Kevin Lin, John Hewitt, Ashwin Paranjape, Michele Bevilacqua, Fabio Petroni, and Percy Liang. Lost in the middle: How language models use long contexts. **Transactions of the Association for Computational Linguistics**, 12:157–173, 2024.
- [5] Razvan Azamfirei, Sapna R. Kudchadkar, and James Fackler. Large language models and the perils of their hallucinations. **Critical Care**, 27(1):120, 2023.
- [6] Yiyang Zhou, Chenhang Cui, Jaehong Yoon, Linjun Zhang, Zhun Deng, Chelsea Finn, Mohit Bansal, and Huaxiu Yao. Analyzing and mitigating object hallucination in large vision-language models. In **The Twelfth International Conference on Learning Representations**, 2024.
- [7] Nils Reimers and Iryna Gurevych. Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In Kentaro Inui, Jing Jiang, Vincent Ng, and Xiaojun Wan, editors, **Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)**, pages 3982–3992, Hong Kong, China, November 2019. Association for Computational Linguistics.
- [8] Jiawei Lin, Jiaqi Guo, Shizhao Sun, Zijiang Yang, Jian-Guang Lou, and Dongmei Zhang. Layoutprompter: Awaken the design ability of large language models. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, **Advances in Neural Information Processing Systems**, volume 36, pages 43852–43879. Curran Associates, Inc., 2023.
- [9] Jianan Li, Jimei Yang, Jianming Zhang, Chang Liu, Christina Wang, and Tingfa Xu. Attribute-conditioned layout GAN for automatic graphic design. **CoRR**, abs/2009.05284, 2020.
- [10] Mayu Otani, Naoto Inoue, Kotaro Kikuchi, and Riku Togashi. Ltsim: Layout transportation-based similarity measure for evaluating layout generation, 2024.

```

""Create content for a research poster from the provided paper. The poster
should be clear, concise, and visually appealing, including the title, author names,
and **4 to 6** main sections summarizing key points. Omit sections like
'Competing Interests,' 'Authors' Contributions,' and 'Acknowledgements.'

- **Title and Authors:** Provide the full title and all authors' names.

- **Content Sections:** Create 4 to 6 sections with:
  - Descriptive titles.
  - Concise summaries of key points, such as methods, findings, and conclusions.
  - Relevant figures, or tables immediately following the related text, using:
    <figure> Figure1: [Title/Description] </figure>, <table> Table3:
    [Title/Description] </table>.
    - Even if it is written in an abbreviation such as fig, please write <figure> and
    "Figure[number]:".
    - Please make sure that the figure [numbers] shown here correspond to the
    numbers within the paper.
  - **Please carefully select and include only up to five figures and tables in your
  poster, focusing on those that are essential and most relevant to your content.**
  - Don't necessarily need to use figures, or tables. Just choose what you need for
  your poster.

- **Style and Format:**
  - Match the content to the paper's abstract and main body.
  - Follow the required format; only summarize the specified sections.

**Output Format:**
Title: [Paper Title]
Authors: [Author Names]
1. [Section Name]
[Section Text]
<figure> Figure1: [Figure Title or Description] </figure>
2. [Section Name]
[Section Text]
<figure> Figure3: [Figure Title or Description] </figure>
3. [Section Name]
[Section Text]
<table> Table3: [Table Title or Description] </table>
4. [Section Name]
[Section Text]
...
""

```

図 4 ポスター生成を目的とした要約文の生成プロンプト

A 付録

A.1 プロンプト

Paper2Poster システムで使用したプロンプトを図 4-6 に示す。これらのプロンプトはタスクの定義や出力形式を詳細に指定して期待する出力を実現している。

A.2 定性的評価指標の詳細

生成した学術ポスターを定性的に評価するための指標の詳細を下記に示す。評価指標は以下の 5 つであり、それぞれ 5 段階 (5 - 非常に良い, 4 - 良い, 3 - 普通, 2 - 悪い, 1 - 非常に悪い) で評価する。

- **視覚性**
観覧者が理解しやすく、見やすいデザインになっているかを評価。
- **構造化**
ポスター内の情報がどれだけ論理的に配置され、良い流れになっているかを評価。
- **明瞭性**
情報が簡潔でありながらも要点が伝わるように構成されているかを評価。
- **網羅性**

```

""Please extract the following data from the summary text:

Title
Author Info
Section
Additionally, for each section (excluding such as "Acknowledgements"), please
organize and provide the following information: Text, Caption, Figure, Table.
If the number of sections is 7 or more, please select and include only the sections
necessary for the poster.
Please place the data in the <> part of output format in the specified format.
**For figures, please include the prefix text, such as Figure1, whenever
possible.**

Format the output as follows:
```json
{
 "Title": <str format>,
 "Author Info": <str format>,
 "Sections": {
 <str format>: {
 "Text": <[list format]>,
 "Figure": <[list format]>,
 "Table": <[list format]>
 }
 }
}
```
""

```

図 5 構造化データの生成プロンプト

```

""Your task is to create a detailed scientific poster layout based on the provided
"section area" information. The layout should be designed according to the
specific sections and their respective ranges, taking into account the content.

**Domain**: Scientific Paper Poster Design

**Canvas Size**: Width: 120px, Height: 160px

**Requirements**:
- Arrange the contents (sections, text, figures, lists) within the designated
"section area" so that they **completely fill the area without any overlap**.
- Use only the provided "section areas" from the Section Area Layout; do not
create new ones.
- Ensure figures are not too wide.
- Position the "Section" title logically at the top of each "section area."
- **Please check carefully whether there are any parts without contents in the
"section area".**
- **Carefully choose the left, top, width and height values.**
- Keep the "Title" and "Author Info" in their designated positions within their
respective section areas.

**Important**:
- The HTML will be parsed automatically, so adhere strictly to the format
provided. Adjust only the number of sections, their sizes, and positions based on
the content volume.
- Provide the layout **strictly** in the following HTML format. Ensure that the
output matches this format exactly, including HTML tags, class names, and inline
styles. Do not include extra whitespace before the <div> tags.

**Example**:

<few-shot example>

Consider the layout while referring to this example.:
""

```

図 6 レイアウトの生成プロンプト

ポスターが論文の重要な情報をどれだけ包括的に提供しているかを評価。

- **正確性**
ポスターに記載されている情報が事実に基づいているかを評価。