

CAMERA³: An Evaluation Dataset for Controllable Ad Text Generation in Japanese

Go Inoue,^{†‡} Akihiko Kato,[‡] Masato Mita,[‡] Ukyo Honda,[‡] Peinan
Zhang[‡]

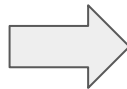
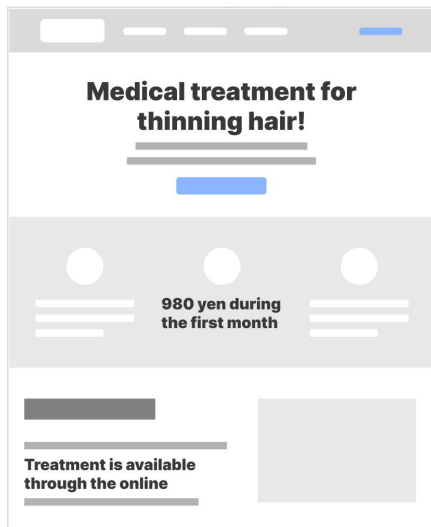
[†]MBZUAI, [‡]CyberAgent



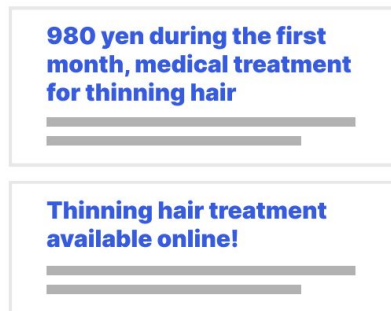
Background

- **Ad text generation** is the task of creating compelling text from an advertising asset, such as Landing Page (LP)

Landing Page (LP)



Ad Text



Motivation

- In advertising, **diversity** plays an important role in enhancing the effectiveness of an ad text
- Diversity in ad text mitigates a phenomenon called “**ad fatigue**”
 - Users become disengaged due to repetitive exposure to the same advertisement
- In this work, we present a dataset where we annotate diverse ad texts for a given LP based on **aspects of ad appeal (A³)**

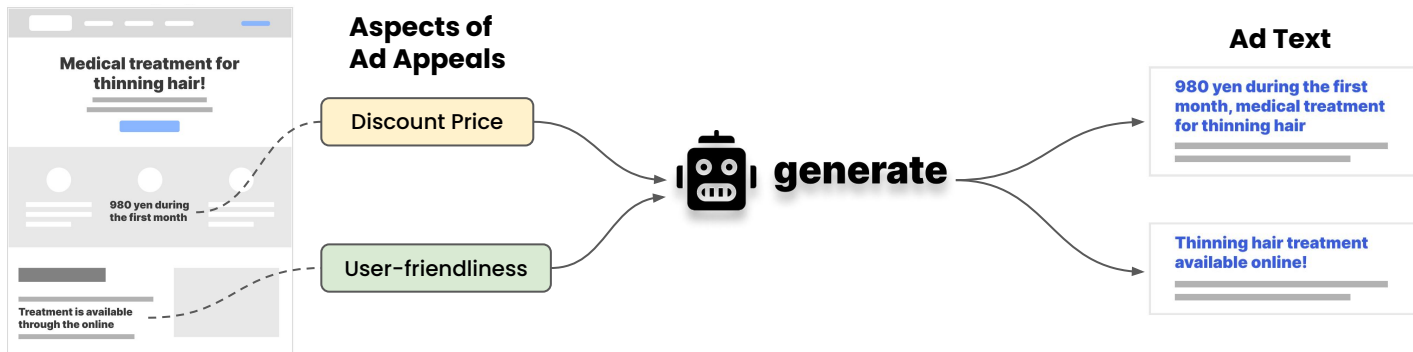
Aspects of ad appeal (A³)

- Ad creators can emphasize various **aspects** when they write ad texts for a certain product
 - e.g.) price, product features, and quality
 - These aspects are called “**aspects of ad appeal (A³)**” (Murakami et al., 2022)
- Example
 - “980 yen during the first month” → emphasizing discount price
 - “Hair treatment available online!” → emphasizing user-friendliness
- In this work, we ask annotators to **identify** these **aspects** in an LP, and to **write an ad text** for each identified aspect

Goal

- **Our goal** is to build/evaluate a model that generates **diverse ad texts** from a single LP, with an ability to **control the aspect** to emphasize
- To that end, we create **CAMERA³**, the **first** evaluation dataset for controllable ad text generation in Japanese

Landing Page (LP)



Related Work

- **CAMERA dataset (Mita et al., 2023)**
 - Four ad texts per LP without aspects of ad appeal (A³) annotation
 - There are multiple ad texts for each LP 👍
 - But we do not know which aspect of ad appeal exists in these ad texts 👎
- **A³ dataset (Murakami et al., 2022)**
 - One ad text per LP with aspects of ad appeal (A³) annotation
 - There is only one ad text for each LP 👎
 - But we know which aspect of ad appeal exists in the ad text 👍
- **This work: CAMERA + A³ = CAMERA³**
 - Multiple ad texts with aspects of ad appeal (A³) annotation

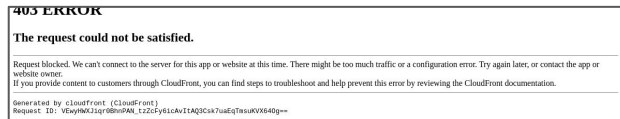
Corpus Creation: Source Data

- **CAMERA dataset (Mita et al., 2023)**
 - The only publicly available dataset in ad text generation in Japanese
 - The test set is a collection of 869 LP images annotated for four ad text per instance
 - Each instance consists of:
 - An LP screenshot image
 - Meta description
 - Keywords
 - Original ad text, and three additionally annotated ad texts
- In this work, we **enrich** the test set of CAMERA dataset with **additional annotations** of aspects of ad appeals and corresponding ad texts

Corpus Creation: Data Preprocessing

- **Manual varification**

- Exclude instances with a screenshot error
- Exclude instances with a significantly lengthy LP image



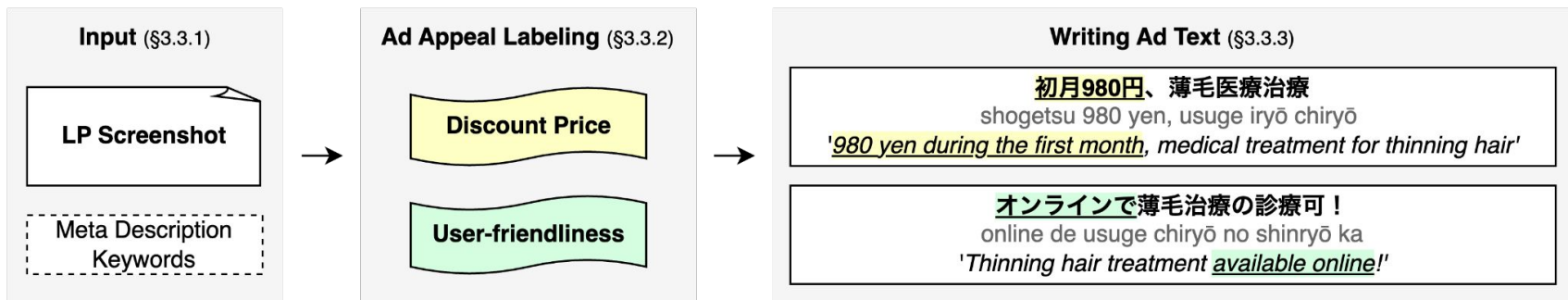
- **Split LP screenshots into four segments**

- For annotation efficiency
- Splitting yields **3,217 LP segments**



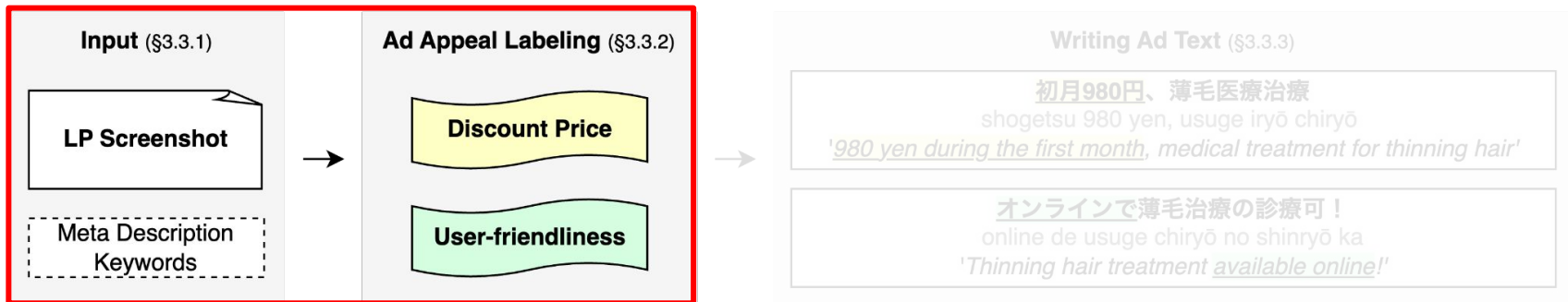
Corpus Creation: Annotation Task

- The annotation task is divided into two subtasks:
 - **Ad appeal labeling** for an LP segment
 - **Ad text creation** for each ad appeal found in the LP segment



Annotation Task: Labeling Aspects of Ad Appeals

- The annotators **identify advertising appeals** in the LP screenshot and annotate them with corresponding aspect labels



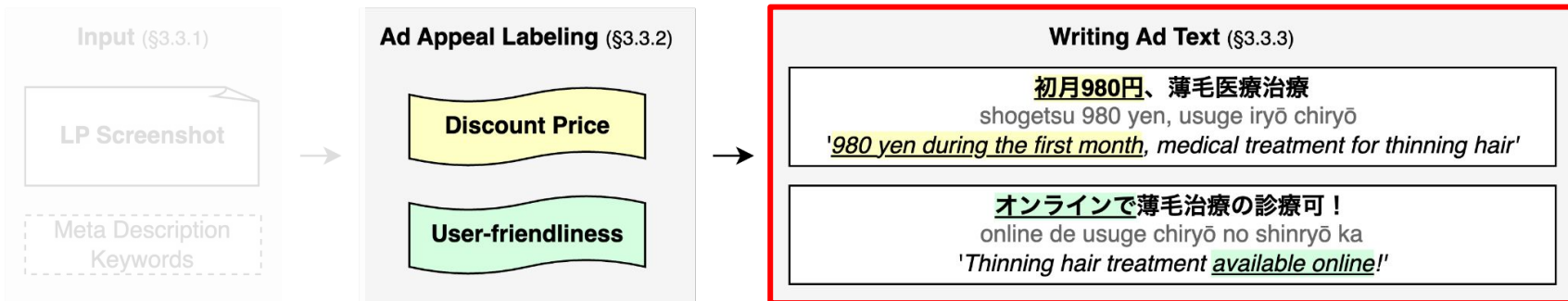
Annotation Task: Labeling Aspects of Ad Appeals

- We use a set of 22 labels based on [Murakami et al. \(2022\)](#)

Ad Appeal Label	
1) No Appeal	12) Other features
2) Discount price	13) Limited time
3) Reward points	14) Limited target
4) Free	15) First-time limited
5) Special gift	16) Other limited offer
6) Other offer	17) Largest/no.1
7) Quality	18) Product lineup
8) Problem Solving	19) Trend
9) Speed	20) Other track record
10) User-friendliness	21) Story
11) Transportation	22) Other

Annotation Task: Writing Ad Text

- The annotators produce **ad text guided by the ad appeal label** assigned during the first stage



Annotation Task: Writing Ad Text

- The ad text **must** follow the following criteria:
 - Ad text must include expressions that represent the annotated ad appeal
 - The length must be within 15 full-width characters (30 half-width characters)
 - Ad text should not be copy-pasted from the LP as is

Annotation Task: Quality Control

- Annotators received training on the annotation process
 - The authors provided explanations of the annotation guidelines
 - Two practice annotation rounds followed by a feedback session by the authors after each round

- Inter-annotator agreement
 - **Krippendorff's alpha** is 0.33 for the labeling task (100 instances)
 - A **fair agreement** based on [Landis and Koch \(1997\)](#)

Corpus Statistics and Format

- **CAMERA³** contains 3,980 annotated ad texts

CAMERA ³ Statistics	
# LP Segments	3,217
# LP Segments w/ Ad Appeal	1,974
# LP Segments w/o Ad Appeal	1,243
# Ad Appeal & Ad Text	3,980

- We release our dataset in json format

```
{
  "instance_id":0,
  "lp_image_sliced":"screen-1200-100303_00.png",
  "annotator_id":5,
  "kw":"My Car, Mutual Insurance",
  "lp_meta_description":"The latest car insurance rankings for 2022!...",
  "lp_image_sliced_ocr_text":"KO insurannce!\nInsurance Comparison KO insurance!...",
  "ad_appeal_type":"Price",
  "ad_text":"Discounts available if you sign up online!"
}
```

*This example is translated for the sake of presentation. The actual data is in Japanese.

Experiment: Ad Text Generation Baseline

- **The goal** is to provide results of a simple baseline as **an initial comparison point** for future efforts
- Input
 - the specified ad appeal label
 - a meta description
 - an OCR-processed LP segment
- We **prompt** a state-of-the-art **LLM** (GPT-3.5-Turbo)
 - Prompt:
 - “You are an advertising copywriter. Write only one advertisement containing {ad appeal label} ({description of the label}) from LP meta description and OCR text. Write it in 15 full-width characters (30 half-width characters) or less. LP meta description: {lp meta description} OCR text: {text}”
 - *We use Japanese for prompting.

Evaluation

- We use the following three criteria:
 - **Content**
 - The content similarity of the generated ad text to the gold reference
 - We use BERTScore (F1) by [Zhang et al. \(2020\)](#)
 - **Style**
 - The percentage of generated ad texts classified to have the target ad appeal label
 - We use a BERT-based classifier by [Murakami et al. \(2022\)](#)
 - **Format**
 - The percentage of generated ad texts following the specified format, i.e., within 15 full-width characters

Results

- The result shows suggests that generating an ad text with a specified ad appeal is a **challenging task**
 - **Low style accuracy** and **very low format accuracy**

GPT-3.5-Turbo	
Content	66.81%
Style	36.31%
Format	13.14%

- This calls for further developments in controllable ad text generation

Conclusion

- We presented **CAMERA³**, an evaluation dataset for controllable ad text generation in Japanese
- Our dataset includes **3,980 ad texts** written by expert annotators, taking into account various **aspects of ad appeals (A³)** within each LP
- Our dataset serves as a testbed for evaluating the capabilities and limitations of state-of-the-art NLG models in **controllable text generation**

Thank you for listening!

- Our dataset is available at <https://github.com/CyberAgentAILab/camera3>



- License: CC BY-NC-SA 4.0