

AP



Text to Image Recommendation

Karina Huang, Dianne Lee, Abhimanyu Vasishth, Phoebe Wong

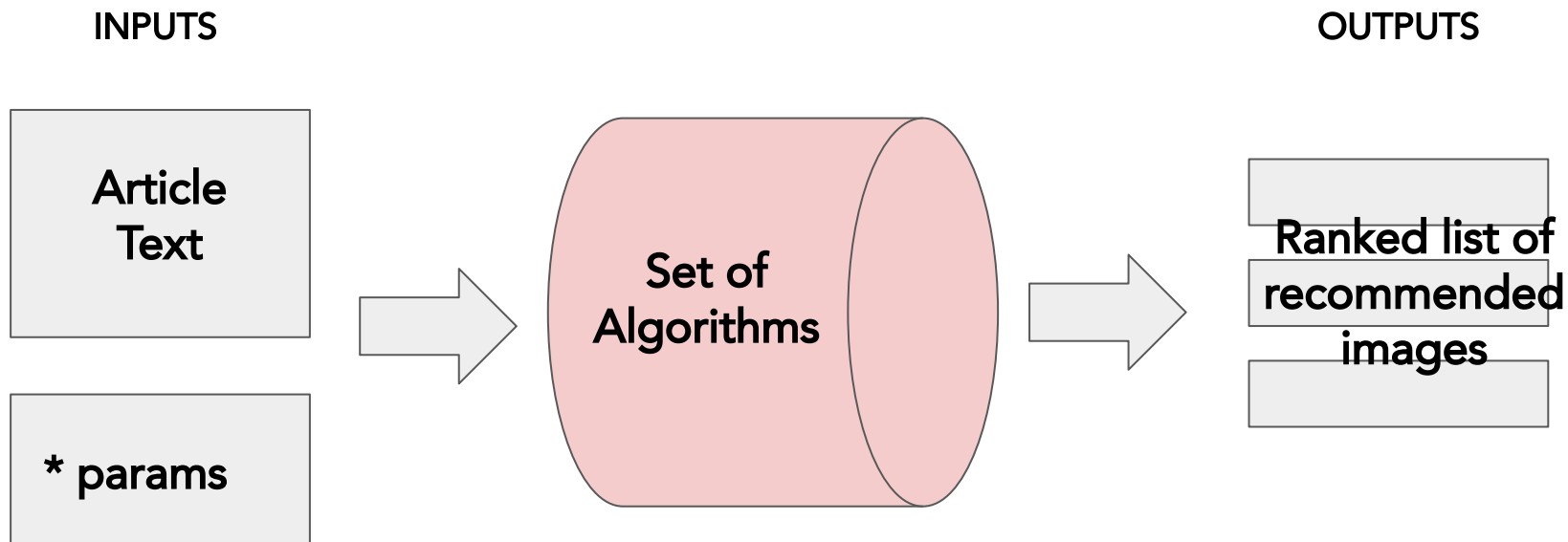


Background

1. What is the Associated Press?
2. What is AP Images?



Project Overview



Use Cases



AP Editors

**Clients
Searching
Articles**

**Clients
Searching
Images**

AP Classification API

- **3-Step Rule-based Tagging Service**

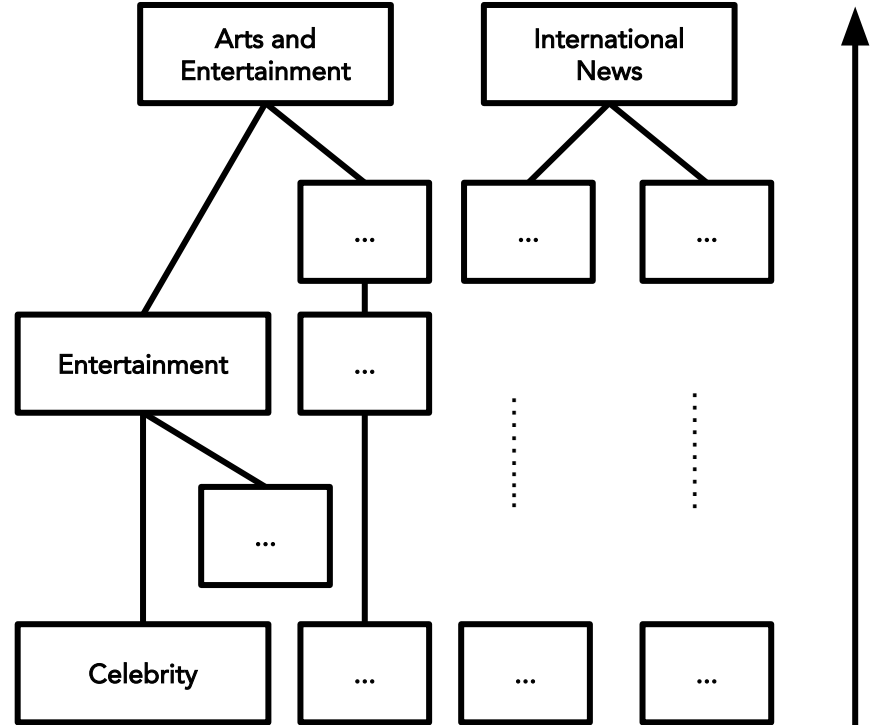
- Input: Text
- Output: Tags
- Sub-Categories -> Categories -> Supplemental Categories

Sam Smith announces new pronouns of 'they' and 'them'

By The Associated Press September 13, 2019



NEW YORK (AP) — Sam Smith has declared the pronouns of “they/them” on social media after coming out as non-binary in what the pop star called a “lifetime of being at war with my gender.”



AP Classification API

- Rule-based Tagging Service

- Input: Text
- Output: Tags
- Sub-Categories -> Categories -> Supplemental Categories

Sam Smith announces new pronouns of 'they' and 'them'

By The Associated Press September 13, 2019



NEW YORK (AP) — Sam Smith has declared the pronouns of “they/them” on social media after coming out as non-binary in what the pop star called a “lifetime of being at war with my gender.”

AP Media API

- Image Search and Retrieval

```
{
  "api_version": "4.1",
  "api_mode": "live",
  "id": "Hy311nqjZ",
  "method": "/content/search.GET",
  "org_name": "My Organization",
  "params": {
    "include": [
      "*"
    ],
    "pricing": true,
    "q": "Emma Stone"
  },
  "data": {
    "query": "Emma Stone",
    "updated": "2017-09-28T17:20:20.160Z",
    "total_items": 3834,
    "current_page": 1,
    "page_size": 10,
    "current_item_count": 10,
    "next_page": "https://api.ap.org/media/v/content/search?page=2",
    "page_template": "https://api.ap.org/media/v/content/search?page={PAGE_NUMBER}",
    "items": [
      {
        "meta": {
          "score": 82.30266,
          "products": [
            {
              "id": 100024,
              "name": "AP Online"
            }
          ]
        },
        "pricing": {
          // Human-readable pricing info
          "policy": {
            // Machine-readable usage instructions and pricing
          }
        }
      },
      // Human-readable usage terms, content metadata and download links
    ]
  }
}
```

AP Content Metadata

- **25,000 data points expected this week**
- **ID**
 - Unique for each revision and update of a piece of content
- **Metadata**
 - Tags / Categories
 - Datetime
 - Content usage information
 - ...
- **Training / Test set**
 - AP News Archive with linked media content (images, and/or videos)
- **Additional Test set**
 - How to model and test for external application?

AP Content Metadata

- **25,000 data points expected this week**
- **ID**
 - Unique for each revision and update of a piece of content
- **Metadata**
 - Tags / Categories
 - Datetime
 - Content usage information
 - ...
- **Training / Test set**
 - AP News Archive with linked media content (images, and/or videos)
- **Additional Test set**
 - How to model and test for external application?

Algorithm/Project First-Step

- Because of existing tags on both text & image → Tag matching
 - e.g., matching “dolphin” tagged text to “dolphin” tagged image

ID:	19184103619156
Creation Date:	July 01, 2019 10:52:51 PM
Submission Date:	July 03, 2019 02:53:04 AM
Photographer:	Stringer
Source:	ICHPL Imaginechina
Credit:	Stringer - Imaginechina
Special Instructions:	For editorial use only. Please contact your AP representative for commercial use. China OUT; France OUT ; IC-URN:bjl11725620
Resolution:	5574 x 3716 4.57 MB
Subject:	Dolphins, Animals
Keywords:	China, Chinese, Guangdong, Zhuhai, pacific, white-sided, dolphin, cub
Location:	Zhuhai, GUANGDONG, CHINA
Transmission Reference:	NYWWP
Byline Title:	CTR
Caption Writer:	Zou Shan
Usage Notes:	This content is intended for editorial use only. For other uses, additional clearances may be required.



Algorithm/Project First-Step

- Because of existing tags on both text & image → Tag matching
 - e.g., matching “dolphin” tagged text to “dolphin” tagged image
- Text: rank tags by relevance/importance
 - Keyword extraction (“are tags representative of the article?”)
 - Theme-weighted PageRank algorithm
 - Thematic weight of a candidate keyword similar to thematic representation of article
 - (nice to have) sentiment analysis of text, then match with sentiment-matching images

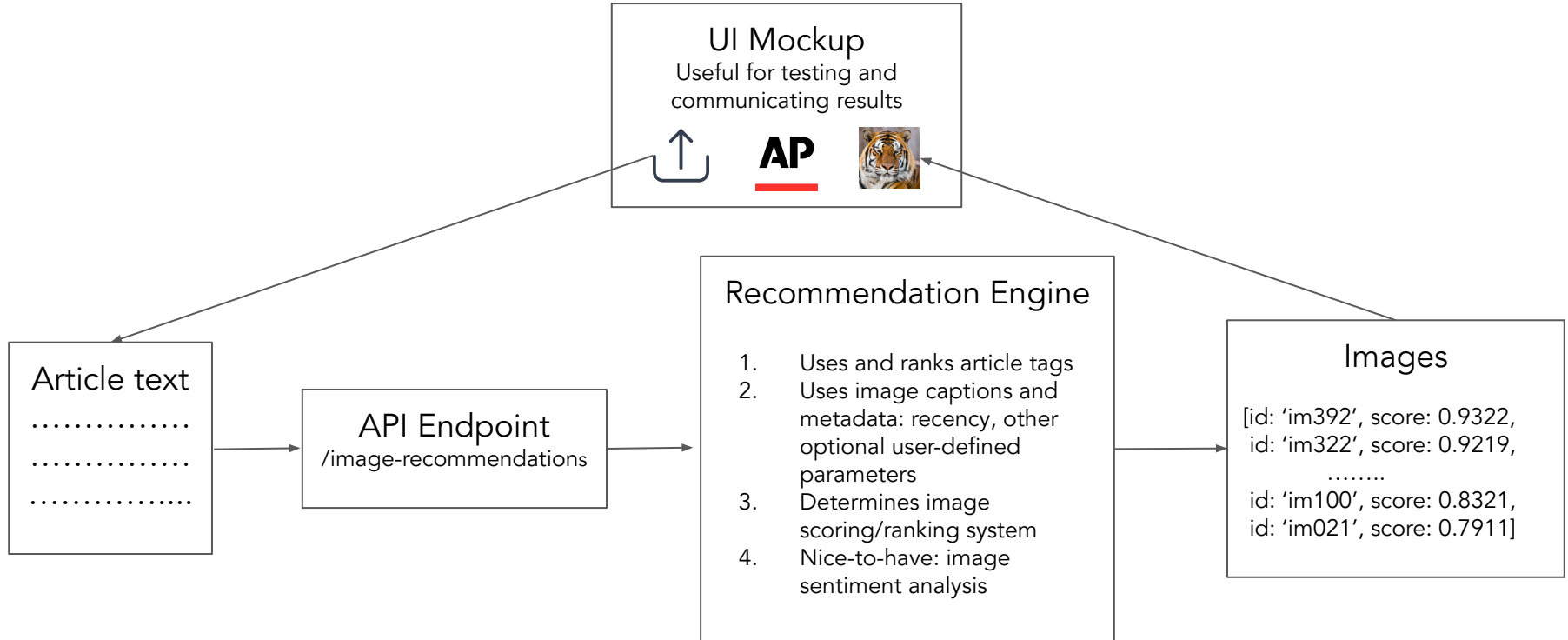
Algorithm/Project First-Step

- Because of existing tags on both text & image → Tag matching
 - e.g., matching “dolphin” tagged text to “dolphin” tagged image
- Text: rank tags by relevance/importance
 - Keyword extraction (“are tags representative of the article?”)
 - Theme-weighted PageRank algorithm
 - Thematic weight of a candidate keyword similar to thematic representation of article
 - (nice to have) sentiment analysis of text, then match with sentiment-matching images

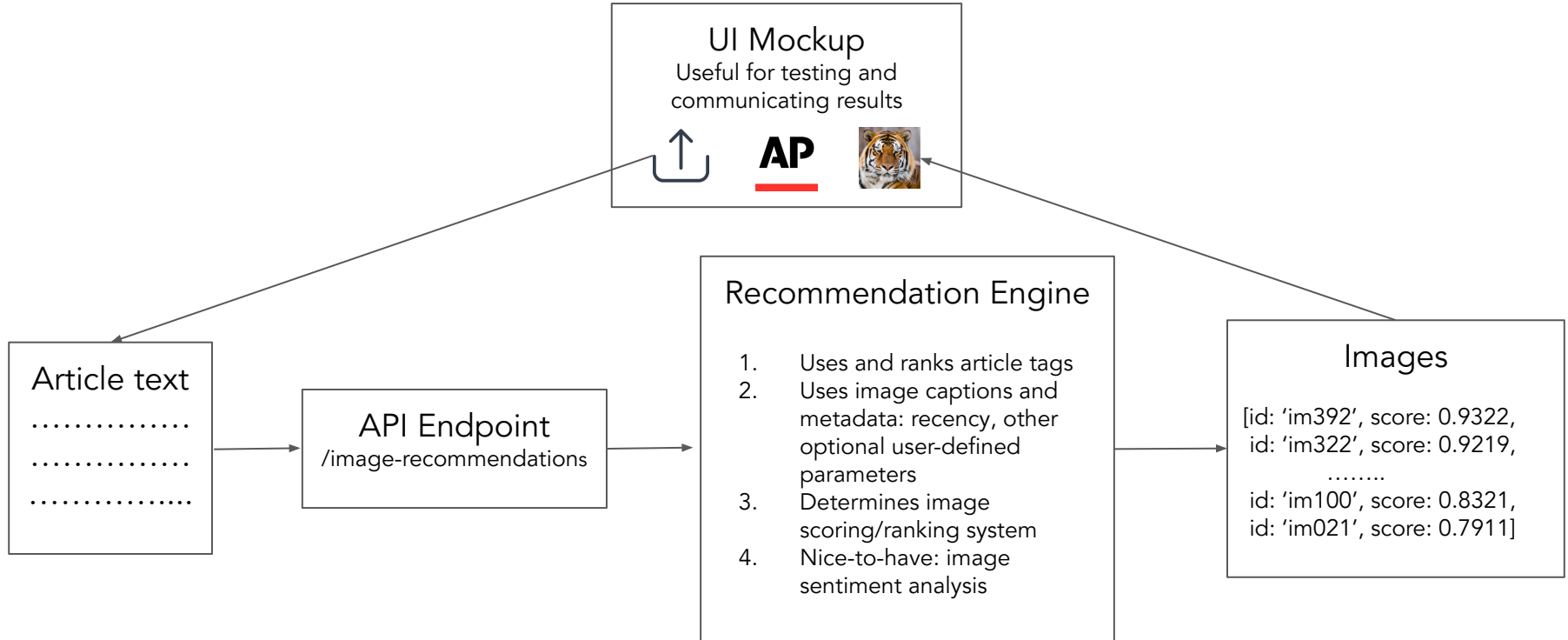
Algorithm/Project First-Step

- Because of existing tags on both text & image → Tag matching
 - e.g., matching “dolphin” tagged text to “dolphin” tagged image
- Text: rank tags by relevance/importance
 - Keyword extraction (“are tags representative of the article?”)
 - Theme-weighted PageRank algorithm
 - Thematic weight of a candidate keyword similar to thematic representation of article
 - (nice to have) sentiment analysis of text, then match with sentiment-matching images
- Image: rank tags by relevance/importance
 - Image classification (probability/area of the tag object in the image)

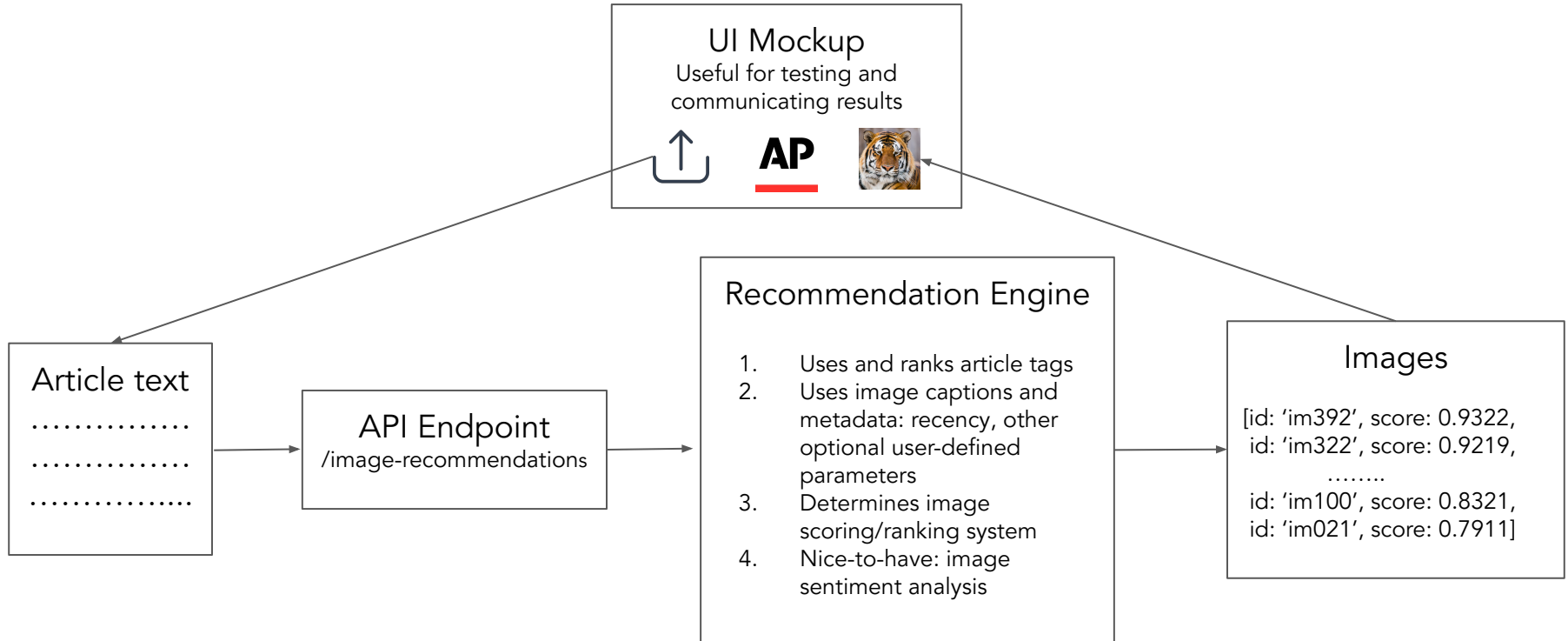
Project Deliverable



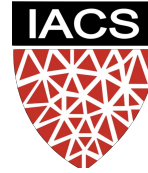
Project Deliverable



Project Deliverable



AP



Thank you! Questions?