Clickbait Spoiling at SemEval 2023

Call for Participation

26th October, 2022

Maik Fröbe\(^1\)
Tim Gollub\(^2\)
Matthias Hagen\(^1\)
Martin Potthast\(^3\)

\(^1\)Friedrich-Schiller-Universität Jena
\(^2\)Bauhaus-Universität Weimar
\(^3\)Leipzig University

@webis_de www.webis.de
Clickbait Spoiling at SemEval’23
Spoilers: The Dark Side
A spoiler is a short summary that reveals details of a plot

Accidentally seeing a spoiler may ruin the experience

Wikipedia had spoiler alerts until 2007
A spoiler is a short summary that reveals details of a plot

Accidentally seeing a spoiler may ruin the experience

Wikipedia had spoiler alerts until 2007
Clickbait Spoiling at SemEval’23
Spoilers: The Light Side

Clickbait [Potthast et al., 2016]

- Social-media posts that create a curiosity gap to generate clicks
- Resolution is usually ordinary or trivial

How to keep your workout clothes from stinking:
lifehac.kr/57YOuEZ
Clickbait Spoiling at SemEval’23
Spoilers: The Light Side

Clickbait  [Potthast et al., 2016]

- Social-media posts that create a curiosity gap to generate clicks
- Resolution is usually ordinary or trivial
Clickbait Spoiling at SemEval’23
Spoilers: The Light Side

Clickbait  [Potthast et al., 2016]

- Social-media posts that create a curiosity gap to generate clicks
- Resolution is usually ordinary or trivial
Clickbait Spoiling at SemEval’23

Spoilers: The Light Side

Clickbait [Potthast et al., 2016]

- Social-media posts that create a curiosity gap to generate clicks
- Resolution is usually ordinary or trivial

Clickbait Spoiling [Hagen et al., 2022]

- Generating a short text that satisfies the curiosity gap
Shared Task on Clickbait Spoiling at SemEval’23

The Dataset

5,000 spoilable clickbait posts with high quality annotations

One weird trick for not losing your iPhone 7 headphone dongle slct.al/2gbEv5

As an iPhone 7 user, I have a better idea:

Just leave the dongle attached to the headphones.

Instead of paying $6 plus shipping for a keychain ring (35 cents) and headphone plug (70 cents), just leave the dongle attached to the headphones you use. That way, when you grab your
5,000 spoilable clickbait posts with high quality annotations

Construction:
- 5 social media accounts that manually spoil clickbait
- Webis Clickbait-17 dataset

560 hours annotation effort:
- Manual post selection, main content extraction, and spoiler identification
- 3 types of spoilers: Phrase, Passage, Multi
Shared Task on Clickbait Spoiling at SemEval’23

Spoiler Types

- **Phrase**
  - \(\leq 5\) consecutive words
  - Often entities and nouns
  
  `Someone is typing.`

- **Passage**
  - \(>5\) consecutive words
  - Often Descriptions
  
  `Just leave the dongle attached to the headphones.`

- **Multipart**
  - Lists closing multiple curiosity gaps
  
  1. Set Your Alarm with *Precision*
  2. Write Down 1 Daily Intention
Shared Task on Clickbait Spoiling at SemEval’23

Task 1: Spoiler type classification

- Input: Post, Article
- Output: Phrase, Passage, Multipart
- Related work:
  - Query intent prediction
    [Alexander et al.: ORCAS-I: Queries Annotated with Intent using Weak Supervision]
  - Query reformulation
  - Query understanding
    [Blog by Daniel Tunkelang on Query Understanding. See https://queryunderstanding.com/]
Shared Task on Clickbait Spoiling at SemEval’23

Task 1: Spoiler type classification

- Input: Post, Article
- Output: Phrase, Passage, Multipart
- Related work:
  - Query intent prediction
    [Alexander et al.: ORCAS-I: Queries Annotated with Intent using Weak Supervision]
  - Query reformulation
  - Query understanding
    [Blog by Daniel Tunkelang on Query Understanding. See https://queryunderstanding.com/]

Task 2: Spoiler generation

- Input: Post, Article, Spoiler type
- Output: text that satisfies the curiosity gap of the post
- Related work:
  - Question answering
  - Passage retrieval
  - Snippet Generation
Shared Task on Clickbait Spoiling at SemEval’23

Connection to your Course

- Apply concepts from the lecture to an unsolved research problem

- Mindset of a Shared Task: [sharki-project.github.io/]
  Friendly scientific competition on a relevant research problem for which its participants develop (computerized) solutions

- Follow your passion. Many directions yield valuable contributions:
  - Effectiveness (e.g., deep learning models or ensembles)
  - Efficiency (e.g., fast removal of non-relevant passages)
  - Creativity (e.g., address the problem in a completely different way)
  - Community impact (e.g., resources that might be helpful for others)
  - Insightfulness (e.g., replace parts of the pipeline with an manual oracle)
  - ...

- Entrypoint to everything: clickbait.webis.de

- Use the forum for questions, help, or if you want to share some resources [tira.io/c/semeval23-task-5-clickbait-spoiling]
Shared Task on Clickbait Spoiling at SemEval’23

Enough Preliminaries...
Shared Task on Clickbait Spoiling at SemEval’23

Enough Preliminaries...

Time to get our hands dirty :)
Shared Task on Clickbait Spoiling at SemEval’23
Software Submissions with TIRA

- We ask you to submit working software
- Benefits: Replicability + Reproducibility
- Procedure:
  1. Implement your approach in Docker images
  2. Upload the Docker image(s) to your dedicated image registry in TIRA
  3. Your approach is executed in a Kubernetes cluster

http://tira.io

TIRA — Evaluation as a Service
Improving the replicability of shared tasks in computer science
Docker Basics

- Goal: You can start/stop/build (your) software
- Docker installation instructions are available for all operating systems [https://docs.docker.com/get-docker/]
- We will provide all commands and resources
Use Cases for Docker

- Run guest systems as containers
- Shipping and running micro services as portable images
- Exploring and experimenting with new technologies
- Encapsulation mechanism to deploy applications in parallel without conflicts

Virtual Machines vs. Docker
Shared Task on Clickbait Spoiling at SemEval’23

Example Docker Commands

- Many images available on hub.docker.com

- Run a bash shell: `docker run --rm -ti bash`
  - `--rm`: Remove container after completion
  - `-ti`: Attach stdin and stdout
  - ToDo: Run above command without `-ti`
  - Attention: Changes within the image are not persistent
Shared Task on Clickbait Spoiling at SemEval’23
Example Docker Commands

- Many images available on hub.docker.com
- Run a bash shell: `docker run --rm -ti bash`
  - `--rm`: Remove container after completion
  - `-ti`: Attach stdin and stdout
  - ToDo: Run above command without `-ti`
  - Attention: Changes within the image are not persistent

- Bash image with volume mounts:
  ```
  docker run --rm -ti -v ${PWD}:/bla bash
  ```
  - `-v <HOST_PATH>:<CONTAINER_PATH>`: Mount the directory specified in `<HOST_PATH>` on the system to the directory specified in `<CONTAINER_PATH>` within the container
Many images available on hub.docker.com

Run a bash shell: `docker run --rm -ti bash`
- `--rm`: Remove container after completion
- `-ti`: Attach stdin and stdout
- ToDo: Run above command without `-ti`
- Attention: Changes within the image are not persistent

Bash image with volume mounts:
```
docker run --rm -ti -v ${PWD}:/bla bash
```
- `-v <HOST_PATH>:<CONTAINER_PATH>`: Mount the directory specified in `<HOST_PATH>` on the system to the directory specified in `<CONTAINER_PATH>` within the container

Similar: Forwarding ports with `-p <HOST_PORT>:<CONTAINER_PORT>`:
```
docker run --rm -ti -p 8888:8888 -v ${PWD}:/home/jovyan/work
jupyter/datascience-notebook
```
Shared Task on Clickbait Spoiling at SemEval’23
Baselines for Task 1: Spoiler Type Classification

- Hands-On: Build, deploy, and run baselines for task 1
  [clickbait.webis.de#baselines]

- Step-by-Step guide is available online
  [clickbait.webis.de#tira-quickstart]

- A screencast going through all the steps is available online
  [www.tira.io/t/screencast-for-participants]
Shared Task on Clickbait Spoiling at SemEval’23

Baselines for Task 1: Spoiler Type Classification

- Hands-On: Build, deploy, and run baselines for task 1
  [clickbait.webis.de#baselines]

- Step-by-Step guide is available online
  [clickbait.webis.de#tira-quickstart]

- A screencast going through all the steps is available online
  [www.tira.io/t/screencast-for-participants]

Naïve Baseline

- Predict always: phrase

Transformer Baseline

- Input: Clickbait post + linked article
- Library: Simple Transformers [https://simpletransformers.ai/]
- Trained in Google Colab [https://colab.research.google.com/]
Baselines for Task 1: Spoiler Type Classification

- Hands-On: Build, deploy, and run baselines for task 1
  [clickbait.webis.de#baselines]
- Step-by-Step guide is available online
  [clickbait.webis.de#tira-quickstart]
- A screencast going through all the steps is available online
  [www.tira.io/t/screencast-for-participants]

Naïve Baseline

- Predict always: `phrase`

Transformer Baseline

- Input: Clickbait post + linked article
- Library: Simple Transformers [https://simpletransformers.ai/]
- Trained in Google Colab [https://colab.research.google.com/]

Effectiveness (Validation Set)

<table>
<thead>
<tr>
<th>Baseline</th>
<th>Balanced Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naïve</td>
<td>33.3</td>
</tr>
<tr>
<td>Transformer</td>
<td>73.4</td>
</tr>
</tbody>
</table>
Shared Task on Clickbait Spoiling at SemEval’23
Suggestions on How to Continue for Task 1

❑ Approaches that we tried that did not work?
  – Feature-based approaches: SVM, Naïve Bayes, XGBoost, LightGBM, ...
❑ Approaches that we tried that worked “better”: Transformer approaches
Shared Task on Clickbait Spoiling at SemEval’23

Suggestions on How to Continue for Task 1

- Approaches that we tried that did not work?
  - Feature-based approaches: SVM, Naïve Bayes, XGBoost, LightGBM, ...
- Approaches that we tried that worked “better”: Transformer approaches

Untested Ideas: Document Engineering

- Identify multipart spoilers via the document structure?
- For instance: Does the document contains an explicit/implicit enumeration?

1. Set Your Alarm with Precision
2. Write Down 1 Daily Intention
Suggestions on How to Continue for Task 1

- Approaches that we tried that did not work?
  - Feature-based approaches: SVM, Naïve Bayes, XGBoost, LightGBM, ...

- Approaches that we tried that worked “better”: Transformer approaches

Untested Ideas: Generate and Test

- Generate phrase, passage, and multipart spoilers
  - Using a wide range of approaches for task 2?

- Use Query Performance Prediction to identify most promising spoiler type

- Would also be very useful in general
  - Only show spoilers when we are confident that they are correct
Shared Task on Clickbait Spoiling at SemEval’23
Suggestions on How to Continue for Task 1

- Approaches that we tried that did not work?
  - Feature-based approaches: SVM, Naïve Bayes, XGBoost, LightGBM, ...
- Approaches that we tried that worked “better”: Transformer approaches

Untested Ideas: Lots of more Ideas :)

- Use pre-trained models from query intent prediction?
  [Alexander et al.: ORCAS-I: Queries Annotated with Intent using Weak Supervision]

- Query reformulation:
  - Is the transformer baseline more effective with clickbait posts reformulated as proper search queries?
  - If so, can the post automatically reformulated?
  - Is proper reformulation all we need? (I.e., spoiling becomes obsolete?)

- Can other query understanding approaches be applied?
  [Blog by Daniel Tunkelang on Query Understanding. See https://queryunderstanding.com/]
Shared Task on Clickbait Spoiling at SemEval’23
That Was a Lot of Content in a Short Time!
Shared Task on Clickbait Spoiling at SemEval’23

Time for Some Rest Before We Continue with Task 2 ...
No... Not this type of rest. But maybe you have some questions?
Shared Task on Clickbait Spoiling at SemEval’23

We Also Have Some Questions

[sharki-project.github.io/]

- What you have learned so far:
  - What a shared task is
  - How to participate in a shared task
Shared Task on Clickbait Spoiling at SemEval’23
We Also Have Some Questions

- What you have learned so far:
  - What a shared task is
  - How to participate in a shared task
- What we would like to learn:

  Are shared tasks suitable for teaching?

- Please participate in a short survey before, during, and after your project
- First survey: umfrage.uni-leipzig.de/index.php/366989
Hands-On: Build, deploy, and run baselines for task 2
[clickbait.webis.de#baselines]

Step-by-Step guide + screencast available online
[clickbait.webis.de#tira-quickstart tira.io/t/screencast-for-participants]
Shared Task on Clickbait Spoiling at SemEval’23
Baselines for Task 2: Spoiler Generation

- Hands-On: Build, deploy, and run baselines for task 2
  [clickbait.webis.de#baselines]

- Step-by-Step guide + screencast available online
  [clickbait.webis.de#tira-quickstart tira.io/t/screencast-for-participants]

Naïve Baseline

- Spoil a clickbait post with the title of the linked article

Transformer Baseline

- Input: Clickbait post + linked article
- Library: Hugging Face Question Answering
  [https://huggingface.co/docs/transformers/tasks/question_answering]
- Pre-trained on squad, later fine-tuned on clickbait dataset in Google Colab
Shared Task on Clickbait Spoiling at SemEval’23

Baselines for Task 2: Spoiler Generation

- Hands-On: Build, deploy, and run baselines for task 2
  [clickbait.webis.de#baselines]
- Step-by-Step guide + screencast available online
  [clickbait.webis.de#tira-quickstart tira.io/t/screencast-for-participants]

Naïve Baseline

- Spoil a clickbait post with the title of the linked article

Transformer Baseline

- Input: Clickbait post + linked article
- Library: Hugging Face Question Answering
  [https://huggingface.co/docs/transformers/tasks/question_answering]
- Pre-trained on squad, later fine-tuned on clickbait dataset in Google Colab

Effectiveness (Validation Set)

<table>
<thead>
<tr>
<th>Baseline</th>
<th>BLEU-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naïve</td>
<td>0.021</td>
</tr>
<tr>
<td>Transformer</td>
<td>0.382</td>
</tr>
</tbody>
</table>
Shared Task on Clickbait Spoiling at SemEval’23
Suggestions on How to Continue for Task 2

- Approaches that we tried that did not work?
  - Passage retrieval / question answering for passage / multipart spoilers

- Approaches that we tried that “worked”:
  - Question answering for phrase spoilers

Ideas: Test more extractive QA approaches for Multipart spoilers

- Advanced IR Course by Sebastian Hofstätter has a tutorial on extractive QA
  [https://www.youtube.com/watch?v=6FNISntK6Sk]
Shared Task on Clickbait Spoiling at SemEval’23
Suggestions on How to Continue for Task 2

❑ Approaches that we tried that did not work?
  – Passage retrieval / question answering for passage / multipart spoilers
❑ Approaches that we tried that “worked”:
  – Question answering for phrase spoilers

Ideas: Recursive Spoiler Generation for Multipart spoilers

❑ Predict an initial spoiler
❑ Remove corresponding spoiler from the linked document
❑ Repeat until complete
Shared Task on Clickbait Spoiling at SemEval’23
Suggestions on How to Continue for Task 2

- Approaches that we tried that did not work?
  - Passage retrieval / question answering for passage / multipart spoilers
- Approaches that we tried that “worked”:
  - Question answering for phrase spoilers

Some more Ideas

- Given a spoiler candidate: predict if the spoiler is complete or not
- Ensemble approaches
- Redo Passage retrieval (did not work for us, maybe we made something wrong?)
- Successively remove non-relevant parts of the document
The clickbait spoiling challenge will be a part of SemEVAL 2023

- Colocated with a major NLP conference

Software submissions to improve reproducibility

- Submit docker images to tira.io

Timeline

- Soon: Register your group
- December 1, 2022: Early bird software submission
- January 10, 2023: Submission deadline
- February 2023 (Optional): Participant paper submission
  
  [Examples: Bondarenko’18, Reusch’21, Zhou’17]
- February 2023 (Optional): Review a foreign paper
- Summer 2023 (Optional): SemEval
Clickbait Spoiling at SemEval’23

The clickbait spoiling challenge will be a part of SemEVAL 2023

- Colocated with a major NLP conference

Software submissions to improve reproducibility

- Submit docker images to tira.io

Timeline

- Soon: Register your group
- December 1, 2022: Early bird software submission
- January 10, 2023: Submission deadline
- February 2023 (Optional): Participant paper submission
  [Examples: Bondarenko’18, Reusch’21, Zhou’17]
- February 2023 (Optional): Review a foreign paper
- Summer 2023 (Optional): SemEval

Thank You!
Clickbait Spoiling at SemEval’23
Ein Paar weitere Lustige Beispiele

- Beispiel aus Deutschland vom 18.10.2022
  [https://twitter.com/niggi/status/1582117029750132737]

- Simpsons Episode zu “Spoiler Boy”
  [https://www.fernsehserien.de/die-simpsons/folgen/31x14-spoiler-boy-1379464]