First Experiences with TIRA for Reproducible Evaluation in Information Retrieval
Tim Gollub, Steven Burrows, Benno Stein

**PAN**

- Competition on plagiarism detection hosted at CLEF.
- Task definition: Given a pair of suspicious and source document, record all passages in the suspicious document that are plagiarized from the source document.
- Evaluation metric is the plagdet score:
  \[ \text{plagdet}(\text{Det}, \text{Truth}) = \frac{F_1(\text{Det}, \text{Truth})}{\log_2(1 + \text{granularity}(\text{Det}, \text{Truth}))} \]

**TIRA**

- Open source experimentation platform for IR research.
- Runs experiments and stores results.
- Disseminates experiments and results as a web service.
- Instances can be connected to a TIRA network.
- Supports organization of competitions.
- Allows public experiments on private data.

### TIRA @ PAN2012

1. **Training Service**

   ![Pan2012/training-service](image1)

   - Detection: detection-42.zip
   - Team: webis
   - Testset: 06_simulated_paraphrase
   - Upload File
   - Search
   - Execute

   ```
tira@model1:~$ unzip -o $Detector -d det &&
   python $PROGRAM/performance.py -p /pan12-training-sets/$Testset/
   -d det > scores.txt
   ```

   - Participants upload detection results for a given training set.
   - From the user inputs the program execution command is generated through substitution.
   - Detection results are unzipped and evaluated with an implementation of plagdet.
   - Participants receive performance results in a result table.
   - The training service served as a leaderboard during the competition.

2. **Evaluation Service**

   ![Pan2012/evaluation-service](image2)

   - Testset: 08_all
   - Detector: kongileilei12
   - Detector: kaspzak12
   - Detector: torrejon12
   - Upload File
   - Search
   - Execute

   - TIRA platforms are provided for two operating systems.
   - Participants submit their plagiarism detection software for deployment on the appropriate instance.
   - A third TIRA instance evaluates all submissions on the private test set and provides the overall results.

**TIRA System Architecture**

![TIRA System Architecture](image3)

- Front-end process
- Back-end process
- HTTP-CLIENT
- TIRA SERVER
- PROGRAM DATABASE
- PROGRAM WRAPPER
- PROGRAM SCHEDULER