

# JJE Evaluation Metric



Funded by  
the European Union



UK Research  
and Innovation

## Usage/License



The software has been licensed under Creative Commons Attribution Non Commercial 4.0 International (CC BY NC 4.0).

## Download



[GitHub repository:](https://github.com/metabolean5/rst-subtree-eval/)  
<https://github.com/metabolean5/rst-subtree-eval/>

## Researchers



- Martial Pastor, Radboud University
- Nelleke Oostdijk, Radboud University  
Centre for Language Studies, The Netherlands

## Summary

The JJE Evaluation Metric is a structure-aware matching algorithm for scoring RST discourse parser output against gold-standard Joint-Joint-Evaluation (JJE) structures extracted from raw, unsegmented social media text. Unlike standard RST evaluation (which assumes gold EDU boundaries), this metric operates directly on full parser output trees, making it suitable for real-world "parsing in the wild" scenarios.

## Design Rationale

- Last-two-EDU anchor: Parser segmentation is imperfect; tail EDUs are the most structurally distinctive.
- Substring rather than exact match: Allows partial EDU splits without penalizing structurally correct predictions.
- Length filter (4-6 EDUs): Constrains search space to JJE definition; prevents spurious matches.
- Inorder pattern extraction: Directly encodes valid JJE tree topologies, avoiding false positives from other triadic RST constellations.

## How It Works

- **Step 1 – Full Tree Parsing:** The parser receives raw text with no pre-identified EDU breaks and produces a complete RST constituency tree.
- **Step 2 – Candidate Extraction:** JJE candidates are extracted via inorder traversal, matching the two structural patterns defined in the paper.
- **Step 3 – EDU Length Filtering:** Candidates filtered to retain only those with 4-6 EDUs.
- **Step 4 – Substring Matching:** The metric compares the last two EDUs of each candidate against the gold structure. Matching on tail EDUs absorbs segmentation errors and maximizes recall without requiring exact span overlap.

## What We Offer

- A standalone Python evaluation script that takes parser output trees (constituency format, compatible with both the Guz & Carenini bottom-up and Liu et al. top-down parsers) and gold JJE annotations, and returns per-class Precision, Recall, and F1 for JJE vs. non-JJE – with no dependency on gold EDU segmentation.



hybrids

HYBRIDS Website



[www.hybridsproject.eu](http://www.hybridsproject.eu)

Get in Touch



[martial.pastor@ru.nl](mailto:martial.pastor@ru.nl)