

# Analyzing the evolution of major categories in Enhancement Proposals in Software Ecosystems

## 1 - Background

Software Enhancement Proposals (SEPs) formalize developer initiatives to software ecosystems

Undocumented rules and the rationale behind this decision process is well studied

Contents of SEPs remain understudied

## 2 - Research Questions

**RQ1** What are the main categories of proposals?

**RQ2** How does the distribution of categories from submitted proposals change over time?

**RQ3** How does the distribution of categories differ between projects?

## 3 - Methodology

Using tags and labels from examined projects, create a universally applicable set of categories

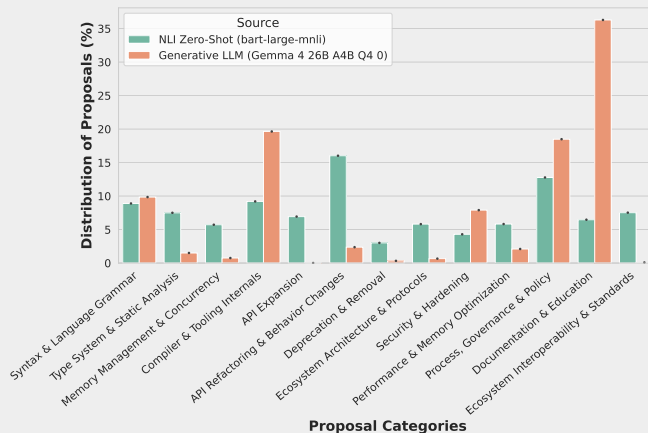
Classify SEPs using Natural Language Inference (bart-large-mnli) and Generative LLM Prompting (Gemma 4)

## 4 - Results

Mitigated two pitfalls in results

- Confounding volume bias
- Semantic overlap bias

### RQ1

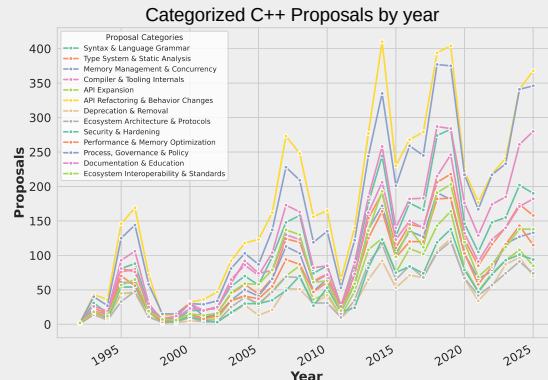


LLM follows semantics; Documentation  
NLI shows intent

## 5 - Conclusions

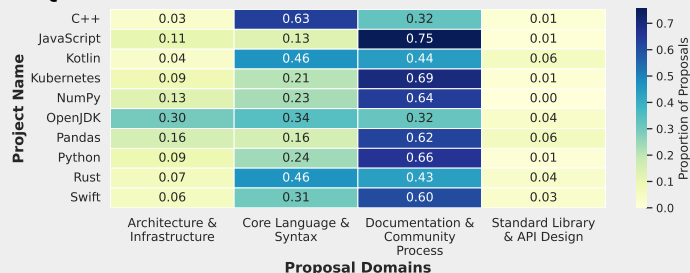
Category distributions remain stable through time  
Planned events are a driving force for development intensity  
Similar projects follow the same patterns

### RQ2



- Relative percentages remain stable
- Absolute volume indicates upcoming events

### RQ3



## 6 - Future Work

(Un)explainable activity spikes may be early indicators for downstream projects  
Capture subtle developer intent with different confidence thresholds