

Beyond the Repo: A Case Study on Open Source Integration with GECToR

Sanjna Kashyap, Zhaoyang Xie, Kenneth Steimel, Nitin Madnani
Educational Testing Service



Task

Integrate the open source **GECToR** code and models (developed by Grammarly) into our production NLP pipeline.

Issues Faced

- Not under active development
- Used older versions of Python, Pytorch and AllenNLP
- No versioning or packaging
- Did not fully exploit AllenNLP's high-level abstractions

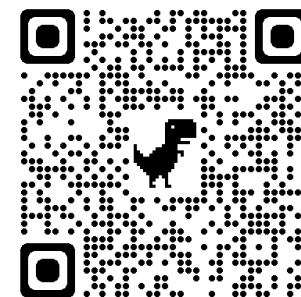
Lessons Learned

Projects should explicitly state a purpose.

Test, test, test!

Estimation of effort is hard but necessary.

Always have a contingency plan.



Link to paper